

**Mechanisms of Transposable Element Repression  
by Piwi Proteins in the piRNA Pathway of *Drosophila* Germ Cells**

Thesis by  
Alexandre Webster

In Partial Fulfillment of the Requirements  
for the Degree of  
Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY  
Pasadena, California  
2015  
(Defended April 28<sup>th</sup>, 2015)

© 2015  
Alexandre Webster  
All Rights Reserved



## ACKNOWLEDGMENTS

The opportunity to present a PhD thesis comes from a cumulative effort across many years and involving the guidance and support of many, many people. I would like to thank all the teachers that have instructed me during my 29 years of education. Among those that stand out are Nancy Wilcox, pre-kindergarten; Dana Skidmore, fourth grade; Arn Krugman, fifth grade; Sarah Barber, sixth grade ancient history; John Fixx, sixth grade English; Terry Hartsoe, sixth grade science; David Acheson, art and geography; Leasile Brickley, music; Armelle Webster, Français (my mother is mentioned twice); Jacob Root, ninth grade ancient literature; Dan Bowen, tenth grade American history; Bill Pickering, eleventh grade western literature; Roger Ipswitch, twelfth grade religious studies; David Harris, general chemistry; Paula Bruice, organic chemistry; Michael Mahan, bacterial pathogenesis; Rolf Christoffersen, genetics; Charles Samuel, virology; Chris Hayes, biochemistry; John Lew, biochemistry; Doug Thrower, oncogenesis and cell cycle; and my undergraduate mentor, Duane Sears, biochemistry and immunology.

My PhD advisor, Alexei Aravin, is the teacher that has spent the most time and effort shaping me into a scientific thinker. I am incredibly fortunate to have had him as an advisor and it is because of him that the years of work in grad school are meaningful. His extraordinary abilities as a PhD mentor made this PhD possible and I am extremely grateful. Thank you Alexei.

Throughout this PhD, Katalin Fejes-Tóth has provided exceptional support in discussing projects, writing papers and practicing talks. Our work together has produced a book chapter that forms the basis of chapter one, in addition to the manuscript presented in chapter two. Thank you so much for your support.

I would like to thank the members of my thesis committee, Paul Sternberg, Bruce Hay, Shu-ou Shan, and David Baltimore, for support, patience, and advice over the years. Your input was invaluable to this PhD.

My greatest teachers also happen to be my parents. My father, Dr. Paul Webster, has supported and encouraged my education from the very beginning. It is in his lab that I first encountered biology and it is because of him that I was able to walk into the world of

science feeling completely at home. My PhD is in many ways because of him and very much for him. My mother, Armelle Webster, has continuously supported and encouraged my education on a daily basis, being my teacher throughout multiple years of school. Her conquests in investigating the Bayeux Tapestry, Impressionism, Concrete poetry among many other subjects made me a curious and well-rounded learner from an early age. Given what she's taught me, I could have just as well written a thesis on the Bayeux Tapestry instead. The encouragement and support of my parents has been unconditional and limitless. Thank you for your endless love and support.

At the most difficult times of this PhD, my friend, brewing partner, and brother Simon Webster was there to encourage me forward. I was very fortunate to have him by my side during the hardships of grad school. It is because of him that I've learned about who I am. We understand our history best. Thank you, Simon.

I am extraordinarily thankful to the AAA Team, first platoon. The friends and colleagues in our lab created a fascinating, fun, and multiculture environment, which was never boring. Ivan Olovnikov, Dubravka Pezic, Irina Meininger, Evelyn Stuwe, Adrien Le Thomas, Ariel Chen, Alicia Rogers, Sergei Manakov, Junho Hur, Hamada Masakazu, Ken Chan, Svetlana Ustyugova, Yicheng Luo, Yixing Cuican (Can Li), Chen-yin Ou, Jae Cho—brothers-and-sisters-in-arms. Thank you for your support, exceptional humor, and for the wildest of times.

I would especially like to thank Evelyn Stuwe for being a great colleague and phenomenal friend, and for being my companion through the highest of highs and the lowest of lows of a PhD. Time travel is possible.

A special thank you goes to my friend and colleague Dubravka Pezic, for the guidance and honesty that I needed at the most critical of times, and for the discussions that put things into perspective.

To Junho Hur, whose scientific wisdom and expertise in biochemistry were essential to this PhD and whose friendship and laughter made late nights/early mornings doing experiments tolerable. Thank you, Junhizzle C3PO.

Thank you to my friend and colleague, Irina Meininger, whose support of our lab is absolutely understated. Thank you for your patience, sense of humor, and for all the wild times.

I would like to thank the students who worked in our lab, especially Xiawei Huang and Elise Cai, who taught me how to teach.

Also, many of the individuals at Caltech that enabled the work of this PhD: Santiago Laparra, for the wisdom of life and encouragement; Andres Collazo, for the expertise that made the microscopy possible; and to Peter Rapp, who introduced me to FRAP techniques.

Lastly, to all thousands of *Drosophila* that shared their gonads for the betterment of scientific knowledge. Bzzzz Bzzzz Bzzzz. You are remembered.

*Per ardua ad astra*

## ABSTRACT

The ability to reproduce is a defining characteristic of all living organisms. During reproduction, the integrity of genetic material transferred from one generation to the next is of utmost importance. Organisms have diverse strategies to ensure the fidelity of genomic information inherited between generations of individuals. In sexually reproducing animals, the piRNA pathway is an RNA-interference (RNAi) mechanism that protects the genomes of germ cells from the replication of ‘selfish’ genetic sequences called transposable elements (TE). When left unabated, the replication of TE sequences can cause gene disruption, double-stranded DNA breaks, and germ cell death that results in sterility of the organism. In *Drosophila*, the piRNA pathway is divided into a cytoplasmic and nuclear branch that involves the functions of three Piwi-clade Argonaute proteins—Piwi, Aubergine (Aub) and Argonaute-3 (Ago3)—which bind piwi-interacting RNA (piRNA) to form the effector complexes that represses deleterious TE sequences.

The work presented in this thesis examines the function and regulation of Piwi proteins in *Drosophila* germ cells. Chapter 1 presents an introduction to piRNA biogenesis and to the essential roles occupied by each Piwi protein in the repression of TE. We discuss the architecture and function of germ granules as the cellular compartments where much of the piRNA pathway operates. In Chapter 2, we present how Piwi in the nucleus co-transcriptionally targets genomic loci expressing TE sequences to direct the deposition of repressive chromatin marks. Chapter 3 examines the cytoplasmic function of the piRNA pathway, where we find that the protein Krimper coordinates Aub and Ago3 in the piRNA ping-pong pathway to adaptively target and destroy TE transcripts. Chapter 4 explores how interactions of Piwis with associated proteins are modulated by arginine methylation modifications. Lastly, in Chapter 5 I present evidence that the cytoplasmic branch of the piRNA pathway can potentially ‘cross-talk’ with the nuclear branch to transfer sequence information to better target and co-transcriptionally silence the genomic loci coding active TE sequences. Overall, the work presented in this thesis constitutes a part of the first steps in understanding the molecular mechanisms that protect germ cells from invasion by TE sequences.

## TABLE OF CONTENTS

<i>ACKNOWLEDGMENTS</i>	<i>iii</i>
<i>ABSTRACT</i>	<i>vi</i>

### ***CHAPTER 1: The piRNA Pathway Guards the Germline Genome Against Transposable Elements***

INTRODUCTION	3
piRNA Biogenesis, Cluster Architecture and Transcription	5
Nuclear and cytoplasmic function of the piRNA pathway	7
Ping-Pong: The Cytoplasmic Function of the piRNA pathway	8
The role of piRNAs in regulating chromatin structure	9
Germ granules: components, structure, and function	11
Nuage are composed of many piRNA pathway components	
in ovarian nurse cells	12
The pole plasm: germ granules in the oocyte	14
Suppressor of Stellate is a specialized piRNA mechanism in Drosophila Testis	14
REFERENCES	16
FIGURES AND FIGURE LEGENDS	28
Figure 1: The Domain Structure of Piwi-Clade Argonaute Proteins.	28
Figure 2: Biogenesis of piRNA in flies begins in the nucleus and initiates the ping-ping cycle in the cytoplasm.	30
Figure 3: Many protein factors make up Nuage granules in Drosophila nurse cells and some accumulate at the pole plasm of the developing oocyte.	32
Table 1: Proteins with identified roles in the Drosophila piRNA pathway.	34

### ***CHAPTER 2: Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state***

### ***CHAPTER 3: Aub and Ago3 are recruited to nuage through two mechanisms to form the ping-pong complex assembled by Krimper***

ABSTRACT	48
INTRODUCTION	48
RESULTS	50
Aub and Ago3 have distinct perinuclear localizations in nuage	50
Aub, but not Ago3 requires a piRNA guide for localization to nuage	52
Ago3 is recruited to nuage by Krimper	53
Tudor domains of Krimper bind the N-terminal region of Ago3	54
Krimper mediates formation of a complex between Aub and Ago3 and is essential for ping-pong piRNA biogenesis	56
DISCUSSION	58
Krimper assembles Ago3 / Aub complexes for ping-pong processing	58
A model for nuage assembly and function in the piRNA pathway	60
ACKNOWLEDGEMENTS	62
MATERIALS AND METHODS	63
REFERENCES	71
FIGURES AND FIGURE LEGENDS	83
Figure 1. Aub and Ago3 have distinct subcellular localizations.	83
Figure 2. Aub requires a piRNA guide for localization to nuage.	85
Figure 3. Krimper recruits Ago3 to granules.	87
Figure 4. Separate domains of Krimper are responsible for dimerization and association with Ago3.	89
Figure 5. Krimper interacts with Aub.	91
Figure 6. Krimper assembles the complex that contains both Aub and Ago3 proteins and is required for ping-pong piRNA processing.	93
Figure 7. A model for recruitment of Aub and Ago3 to nuage and formation of the ping-pong piRNA processing (4P) complex.	95
SUPPLEMENTAL FIGURES AND FIGURE LEGENDS	97
Supplemental Figure 1: The expression of Aub and Ago3 transgenes in Drosophila ovaries.	97
Supplemental Figure 2: Dynamics of Aub and Ago3 in nuage.	99

Supplemental Figure 3: Krimper multimerization is mediated by self-interactions in the N-terminal region.	101
Supplemental Figure 4: Ago3 does not co-immunopurify with different nuage components in ovaries; Granularization of Ago3 fragments in S2 cells.	103
Supplemental Figure 5: The effect of Krimp, Aub, Ago3, and Piwi mutations on ping-pong piRNA processing.	105
Supplemental Figure 6. Evaluating the specificity and efficacy of knockdown fly strains.	107
 <b>CHAPTER 4:</b>	
<b><i>Arginine methylation signals the piRNA-loaded states of Aub and Ago3 to enable interactions with Tudor domain proteins</i></b>	<b>109</b>
INTRODUCTION	110
RESULTS	112
The N-terminal region is the most diverse region of Piwi proteins but is not essential for piRNA loading	112
N-terminal arginine methylation is not required for localization or piRNA loading of Aub or Ago3	115
Aub arginine methylation is essential for progression of the piRNA ping-pong pathway and fertility	117
piRNA loading is required for arginine methylation	118
DISCUSSION	120
Arginine methylation of Aub and Ago3 N-terminal residues is largely dispensable for their recruitment to nuage and for loading with piRNAs	120
Arginine methylation occurs downstream of piRNA loading	122
REFERENCES	125
MATERIALS AND METHODS	137
FIGURES AND FIGURE LEGENDS	142
Figure 1: Identity and Divergence analysis of Drosophila Piwi-clade Argonaute proteins.	142
Figure 2: N-terminal domain swaps reveal localization and piRNA loading of Piwi, Aub, and Ago3 are dependent on PAZ-mid-PIWI domains.	144

Figure 3: Reciprocal N-terminal domain swaps between Ago1 and Aub confirm the specificity of PAZ-mid-PIWI domains in piRNA and miRNA loading and localization.	146
Figure 3. Arginine methylation has a minor role in Aub or Ago3 nuage localization, mobility and the loading with piRNAs.	148
Figure 5 Aub localizes to nuage even in the absence of PRMT5 methylosome components, Csu1 and Vls.	151
Figure 6: Arginine methylation of Aub is required for fertility and interactions with Krimp.	153
Figure 7: piRNA binding-deficient mutant Aub protein, Aub-YK, does not localize to the pole plasm.	155
Figure 8: A model for arginine methylation regulating Aub and Ago3 in the piRNA pathway.	157

## **CHAPTER 5:**

### ***Aub endonucleolytic function stabilizes an intermediate complex containing Piwi.***

	159
INTRODUCTION	160
RESULTS	161
Endonuclease activity is required for the mobility of Aub and Ago3 in nuage	161
Impairment of Aub endonuclease activity stabilizes its interaction with SpnE and Mael	162
Piwi devoid of piRNA forms a complex with Aub	163
DISCUSSION	163
piRNA directs Aub recruitment to nuage, while Aub nuclease activity releases it from nuage	163
Aub interacts with piRNA pathway components, SpnE, and Mael in a pre-cleavage complex that recruits Piwi	164
The possibility that Aub directs the loading of piRNA into Piwi	165
REFERENCES	167
MATERIALS AND METHODS	180
FIGURES AND FIGURE LEGENDS	184



Figure 1. Endonuclease activity is required for the mobility of Aub and Ago3 in nuage.	184
Figure 2. Catalytic activity is required for the mobility of Aub and Ago3, but does not affect pole plasm localization or loading of piRNA into Aub.	186
Figure 3. Co-localization of catalytically dead Aub (Aub-CD) with other nuage components.	188
Figure 4. Catalytic inactivation of Aub leads to stabilization of its interactions with SpnE and Mael.	190
Figure 5. The interaction between Aub and Piwi is stabilized by nuclease-deficiency of Aub and piRNA-binding deficiency of Piwi.	192
Figure 6. A model for assembly of Aub pre-cleavage complex that are recruit Piwi.	194
Supplemental Figure 1. Mean Pearson's correlation coefficient of reciprocal pairs of GFP-tagged or mK2-tagged Aub-WT, Aub-CD, and Ago3-WT proteins co-expressed in nurse cells.	196

## CHAPTER 1

### ***The piRNA Pathway Guards the Germline Genome Against Transposable Elements***

Katalin Fejes Tóth\*, Dubravka Pezic\*, Evelyn Stuwe\* & Alexandre Webster\*

\* Authors contributed equally to this work.

This chapter contains parts a book chapter that is in press:

Non-coding RNAs and the reproductive system (ed. Dagmar Wilhelm), 2015.  
*Springer Advances in Experimental Medicine and Biology*

## INTRODUCTION

Of the 100 trillion cells in the body, germ cells are the only ones contributing to all future generations, and as such are essentially immortal. It is therefore of pivotal importance to preserve the integrity of the germ cell genome, as damage to it can result in an evolutionary dead end—sterility.

Transposable elements (TEs)—highly abundant and repetitive segments of DNA with the ability to replicate and insert into new genomic locations wreaking havoc in the process—pose one of the greatest internal threats to genome integrity. New insertions cause breaks in the DNA, which lead to cell cycle arrest, disrupt genes and their regulatory regions, and increase the likelihood of non-homologous recombination [1]. Various systems have evolved in different organisms to combat and contain the expansion of TEs, but it is the elegant mechanism employed by the germ cells of metazoans that has attracted considerable interest in recent years: the piRNA pathway [2].

The piRNA pathway is a germ line-specific RNA silencing mechanism. The central effector complex of the pathway, called pi-RISC (piRNA-induced silencing complex) in analogy to canonical RNA interference pathways, consists of a protein from the Piwi subfamily of Argonaute nucleases, and a Piwi-interacting RNA, or piRNA. The piRNA pathway relies on the effector function of the Argonaute protein and the specificity provided by the piRNAs to restrict TE activity. Piwi proteins compose a clade of the Argonaute protein family that is specific to metazoans and shows gonad-specific expression [3-8]. Piwi—the protein that is the namesake of the Piwi clade of Argonautes—was discovered in a screen for factors affecting stem cell maintenance in the *Drosophila* germ line. Mutants of this gene have very small gonads, and were therefore named *P*-element induced *w*impy testis, or *piwi* mutants [4, 7].

The domain structure of Piwi proteins is similar to that of Argonautes (**Figure 1**) [9, 10]. Like Argonautes, the mid domain anchors the 5' end of a piRNA, and the PAZ domain lodges the 3' end of the piRNA. The PIWI domain contains catalytic residues within an RNase H-like nucleolytic fold, providing catalytic cleavage of target transcripts. Unlike Argonautes, however, Piwis have Arginine-rich motifs near their N-termini. These residues are post-translationally dimethylated. Arginine methylation allows Piwi

proteins to interact with components of the pathway from the Tudor family [11-14]. The eponymous Tudor domain can bind methylated arginines on Piwi proteins, and this relationship appears to be conserved in many species. Many Tudor proteins have several Tudor domains and act as a scaffold for the formation of higher-order complexes [15, 16]. In germ cells of *Drosophila* and in mouse testis, this mode of protein interaction translates into formation of distinct perinuclear granules similar to p-bodies, which harbor components of the piRNA pathway and have been known for a long time in cell biology as “nuage” [17-20].

Based on shared domain structure with Argonaute proteins, it was inferred that Piwi proteins must also bind small RNAs. Indeed, with the emergence of high throughput sequencing techniques Piwi-associated small RNAs from several species were sequenced and characterized [21-24]. piRNAs are typically 24-30 nucleotides (nt) in length and are longer relatives of ubiquitous short interfering RNAs (siRNAs) and micro RNAs (miRNAs). Based on their sequence, piRNAs and the conserved role of Piwi proteins in taming transposable elements in the germline of animals were elucidated.

piRNAs are generated either from RNA transcripts of active TE copies or from transcripts originating from specialized loci in the genome called piRNA clusters. Clusters are loci harboring mostly defunct remnants of TEs and form the basis of immunity against TE propagation [17, 19, 25]. The piRNAs that are generated from piRNA clusters are mostly antisense to TE mRNA sequences and serve as guides for Piwi proteins to find TE transcripts by complementary base pairing. In the cytoplasm, the identification of target mRNAs by Piwis leads to cleavage of TE transcripts resulting in destruction of the TE message and the concomitant amplification of defensive sequences targeting active TEs as the cleavage product itself is processed into piRNA. The piRNA pathway has often been compared to an adaptive immune system, as it conveys memory of previous transposon invasions by storing TE sequence information in piRNA clusters. By amplifying piRNAs that are complementary to the active transposon sequence, the piRNA pathway can respond quickly and specifically to acute TE activation. In addition to targeted cleavage of TE mRNAs within the cytoplasm, Piwi proteins also function on the chromatin level to silence TE transcription by guiding formation of repressive

heterochromatin on active copies of TEs [26, 27]. Thus, the piRNA pathway acts on two levels to limit germline transposable element activity.

### ***piRNA Biogenesis, Cluster Architecture and Transcription***

piRNAs are classified into two groups based on their biogenesis: primary and secondary piRNA. Primary piRNA are generated by a currently poorly characterized biogenesis machinery, while secondary piRNA arise in an amplification mechanism termed the Ping-Pong amplification loop to specifically enhance piRNA sequences targeting active elements. In this section we will concentrate on what is known about biogenesis of primary piRNAs, focusing on their genomic origin, regulation of transcription, and factors involved in their processing.

piRNA populations are highly complex: deep sequencing of piRNAs from mouse and fly revealed millions of individual, distinct piRNA molecules. Neither piRNAs nor their precursor sequences show any structural motif or sequence bias except for a preference for a Uracil as the first 5' nucleotide (1U bias) of the piRNA. However, when mapped to the genome, this highly diverse population of piRNAs can be mapped to a few discrete genomic loci, called piRNA clusters [19, 21-25, 28]. These clusters are transcriptional units up to 200 kilobases long that— at least in flies— are mostly located in pericentromeric and subtelomeric beta-heterochromatin and are highly enriched in transposable element sequences [19]. In flies, primary piRNAs are generated almost exclusively from these piRNA clusters. It is believed that upon new transposon invasion the TE eventually “jumps” into one of the clusters, leaving a memory of the invasion. Once a transposable element leaves a trace in a cluster, it can be targeted by the piRNA machinery, which will limit further propagation of the TE. Currently it is not known what causes active transposons to preferentially integrate into clusters, but the chromatin features of clusters have been proposed to play a role.

In fly ovaries, clusters are expressed in two cell types: cells of germline origin that include the developing oocyte and associated nurse cells in addition to somatic support cells called follicular cells. Interestingly, the structure of clusters seems to differ depending on where they are expressed: germline clusters are transcribed bi-directionally, generating both sense and antisense piRNA in relation to the transposon mRNA.

Conversely, somatic clusters appear to be transcribed uni-directionally, producing mostly piRNA that are antisense to TE coding regions [19]. Overall, the source from where a piRNA is transcribed dictates how it is processed.

piRNAs in flies are 23-25nt in length, whereas in mouse they are slightly longer, averaging 25-28nt. Similar to other small RNA classes, piRNAs are processed from longer precursors; however, in contrast to miRNAs and siRNAs, the precursors are single stranded transcripts without obvious hairpin structures. It is thought that piRNA biogenesis begins with the endonucleolytic cleavage of the long precursor transcript, generating shorter piRNA precursors. This cleavage event possibly specifies the 5' end of the future piRNA. Genetic screens and structural studies suggest that the endonuclease, Zucchini (Zuc), might conduct this first cleavage of some piRNA precursors (Figure 2A) [29-35]. After cleavage, the 5' end of piRNA precursors gets loaded into a Piwi protein. In flies, primary piRNA are loaded into two of the three Piwi proteins, Piwi and Aubergine (Aub). The factors that make up the piRNA loading machinery are currently unknown; however, several studies have identified the proteins Shutdown (Shu), Vreteno (Vret), Brother-of-Yb (BoYb), and Sister-of-Yb (SoYb) as components that are necessary for loading of Piwi and Aub with primary piRNA[30, 32] (Table 1). Loading of Piwi seems to require some additional factors including Armitage (Armi) and Zucchini (Zuc) and in somatic cells, the Tudor domain protein, Yb[32, 36].

Primary piRNAs show a strong bias for a uridine as their 5' terminal nucleotide. At which step this bias is introduced is not clear and there are two conceivable scenarios. First, cleavage of precursor transcripts could be random and the 5' binding pocket of Piwi proteins could have a steric preference for binding piRNAs with a 5' terminal 1U; or second, the nuclease conducting the first cleavage on cluster transcript precursors could preferably cleave upstream of uridines, thereby creating the observed bias.

Once loaded into Piwi proteins, the piRNA precursor is trimmed at its 3' end by an unknown 3' -> 5' exonuclease. This nuclease is tentatively named "Trimmer" (Figure 2B). The final length of the piRNA is determined by the piRNA binding pocket of the Piwi protein: Piwi associated piRNAs are 25 nt long, whereas piRNAs associated with AUB or AGO3 are 24 nt or 23 nt long, respectively. However, piRNA can vary in length

by one to two nucleotides, possibly as a consequence of imprecise trimming. Finally, after trimming, piRNAs acquire a characteristic 2' methyl modification at their 3' end that is introduced by the piRNA methyl-transferase Hen-1 and leads to stabilization of the small RNA [37].

Overall, while the exact biochemistry and the involved factors are still not fully known, piRNA biogenesis consists of multiple consecutive steps from transcription of precursors, through precursor cleavage and loading, to trimming and 2' O-methylation. Defects in any of these steps leads to diminished piRNAs, upregulation of transposons, and sterility.

### ***Nuclear and cytoplasmic function of the piRNA pathway***

Transposons can be classified into DNA or retrotransposons based on the mechanism by which they transpose in the genome. DNA transposons mobilize by a 'cut and paste' mechanism that is accomplished with the help of an encoded transposase protein. Retrotransposons, further divided into long terminal repeat (LTR) and long interspersed nuclear element (LINE) classes, propagate by reverse transcription of an RNA intermediate. Here too, the required enzymes are encoded by the transposable elements. Another class of TE is the semi-autonomous sequences, such as Alu repeats, that require the protein machinery of other transposons for their replication. In either case the transposable element needs to be transcribed as an mRNA because it encodes the enzymes necessary for transposition, or is also required as a template for reverse transcription. Accordingly, transposition can be regulated on multiple levels, from the efficiency of transcription to transcript stability and rate of translation of required factors, or by interfering with the process of re-integration into the genome. The piRNA pathway targets two steps that are required for all transposons: in the cytoplasm, the piRNA pathway directly targets and destroys RNAs of transposons that escaped transcriptional silencing. In the nucleus, piRNAs are implicated in the regulation of chromatin structure that can affect transcription of targeted loci. In this section we will describe the details of these two levels of defense against transposable elements.

### ***Ping-Pong: The Cytoplasmic Function of the piRNA pathway***

The beauty and power of the piRNA pathway lies in its ability to store information of previous TE invasions (in piRNA clusters) and to specifically respond to TE activation by distinctively amplifying sequences complementary to active elements. The latter is achieved through the interplay of two cytoplasmic Piwi proteins— in *Drosophila*, Aubergine and Argonaute 3— in a process termed the Ping-Pong cycle. As these two proteins are only expressed in the germline, the Ping-Pong cycle is also restricted to germ cells [19].

In *Drosophila*, Aubergine (Aub) and Argonaute-3 (Ago3) act as partners in the defense against active transposable elements, with Aub binding to piRNAs mainly coming from piRNA cluster transcripts and Ago3 enriched in piRNAs from transposon mRNAs [19, 38]. Unlike Piwi, both Aub and Ago3 can catalytically cleave RNA targets [39]. At the beginning of the pathway, Aub loads with primary piRNA sequences that are derived from piRNA clusters and are antisense to TE transcripts. This primary piRNA guides Aub to scavenge the cytoplasm for complementary transcripts. When a TE mRNA with sequence complementarity to an Aub-bound piRNA is identified, Aub cleaves the target RNA ten nucleotides downstream of the piRNA 5' end. This cleavage of transposon mRNA serves two purposes: first, it destroys the transposon transcript and second, it generates the 5' end of a new piRNA precursor (Figure 2C). This new precursor gets incorporated into Ago3 and is processed into a so-called secondary piRNA (Figure 2D). Due to the fact that Ago3-bound piRNA is in sense orientation to the transposable element, this piRNA can promote the production of a new round of cluster-derived piRNA by recognizing complementary cluster transcripts (which then get cleaved and incorporated into Aub) (Figure 2E). As the cycle can only function if transposon mRNA is present, the Ping-Pong cycle shapes a specific response against active TEs.

Since Aub-bound piRNAs have a strong bias for uracil at their 5' ends, the piRNAs that are loaded into Ago3 have a bias for Adenosine at position 10 (10A bias) and are complementary to the Aub-bound piRNAs over a 10 nucleotide stretch. This relationship between Aub piRNA and Ago3 piRNA is termed the Ping-Pong signature [19]



The precise mechanism of how Aub and Ago3 are coordinated in the Ping-pong cycle is not yet fully understood. Protein components that are implicated in the Ping-pong cycle include not only Aub and Ago3, but also the Tudor domain-containing proteins Qin, Krimper (Krimp), and Spindle-E (SpnE) in addition to the essential germline helicase, Vasa (SpnE and Vasa are both putative DEAD-box RNA helicases, and Qin contains a RING E3 ligase domain with unknown function) [20, 40-43]. In this thesis, we present a model for Ping-pong that identifies Krimper as the essential molecular factor necessary to bridge Aub and Ago3 during ping-pong. The model proposes that a dimer of Krimper provides one binding site for piRNA-unloaded Ago3 protein, and a second binding site that can accommodate methylated arginines of piRNA-loaded Aub protein. The model is explained in greater depth in Chapter 3.

### ***The role of piRNAs in regulating chromatin structure***

Transcriptional regulation of specific loci is influenced by the overall chromatin architecture of its genomic environment. In mammals, methylation of DNA and various modifications of histone proteins, as well as histone variants, play a role in this regulation. In *Drosophila melanogaster*, histone modifications and histone variants are mainly responsible for defining the properties of chromatin. In both flies and mice, one of the three Piwi proteins localizes to the nucleus, suggesting a nuclear role for the piRNA pathway, possibly through transcriptional repression of targets [4, 17]. Such a mechanism is further strengthened by the observation that in flies Piwi shows a banding pattern on polytene chromosomes of nurse cells, suggesting a direct interaction with chromatin [26].

In fruit flies the piRNA pathway induces transcriptional repression by initiating deposition of repressive histone marks on target loci. Indeed, multiple studies have shown that the piRNA pathway transcriptionally represses at least some transposons. In cell culture experiments, deletion of Piwi's nuclear localization signal leads to failure of transposon silencing, even though Piwi proteins are loaded with piRNAs in the cytoplasm [44]. Similarly, in flies the deletion of the N-terminus of Piwi leads to its delocalization from the nucleus and a slight increase in active histone marks (di-methylation of H3K79 and H3K4) and a decrease of repressive marks (di-/tri-methylation of H3K9) over several transposons [45]. Additionally, an increase in transcription and accumulation of several telomeric retrotransposons was observed in *Drosophila* ovaries upon mutation of piRNA

pathway components [46-48]. The telomeric transposons, which showed increased transcription, showed slight changes in their chromatin marks [48]. These observations all indirectly hinted towards an involvement of Piwi in transcriptional regulation of transposable elements. Recently, three genome-wide studies confirmed this assumption: knockdown of Piwi leads to transcriptional derepression of a significant fraction of TEs both in cell culture and in ovaries as assessed by Pol II ChIP-seq and by GRO-seq [26, 27, 49]. Upon Piwi knockdown transcript levels of transposable elements increase significantly, indicating that Piwi primarily— if not solely— represses transposons at the transcriptional level.

While it seems likely that transcriptional silencing of transposable elements is achieved through establishment of a repressive chromatin state, this correlation between repressive chromatin marks and silencing has not yet been shown directly. Likewise, factors involved in mediating Piwi's repressive function are not known. Recently, two factors have been proposed to be involved in inducing transcriptional silencing. Maelstom (Mael) is a conserved factor that has been implicated in TE silencing in both mouse and flies [18, 27, 50]. Knockdown of Mael in flies leads to increased Pol II occupancy over TEs similar to changes observed upon Piwi knockdown [27]. Interestingly, while Piwi knockdown resulted in strong reduction of H3K9me3 signal at and downstream of target sequences, Mael knockdown resulted in increased spreading and only a modest reduction in H3K9me3 signal, indicating that Mael does not act in establishing the repressive H3K9me3 mark over targets. Another recent study identified Asterix (CG3893) as a factor required for Piwi-mediated establishment of H3K9me3 mark over at least a subset of TEs [31].

A study using a cell-culture model showed that Piwi silences a significant number of novel TE insertions through tri-methylation of H3K9 [27]. This silencing mark does not only cover the site of the novel TE insertion, but spreads up to 15 kb downstream in the direction of transcription. This observation indicates that transcription plays a role in establishing repressive chromatin. The spreading of repressive marks upon insertion of a transposable element can even lead to repression of proximal protein coding genes. In fly ovaries, however, similar new insertions in proximity to protein coding genes were not observed [26]. It is conceivable that while in cultured cells the silencing of nearby genes

is tolerated and can be detected, in a living organism there is a strong selective pressure against insertions that would compromise functional gene expression.

The piRNA sequences bound to Piwi identify the sequences that are targeted for Piwi-mediated transcriptional repression. This is in agreement with the observation that host genes are generally unaffected by the pathway and that Piwi requires piRNAs to repress TEs [26]. Additionally, loading Piwi with artificial sequences mapping to the lacZ transgene leads to silencing of lacZ-expressing loci in vivo. Additionally, the target locus acquires H3K9me3 marks, which in turn seems to attract further components of the piRNA pathway [26]. While the mechanism of target recognition remains unresolved, it is likely that piRNAs recognize nascent transcript targets through sequence complementarity. This is supported by the observation that Piwi knockdown results in decreased H3K9 tri-methylation of the genomic environment of active transposons, while untranscribed fragments of transposable elements throughout the genome remained unaffected [27].

Transcriptional repression by Piwi occurs in both germ cells and follicular cells of *Drosophila* ovaries. Follicular cells are the somatic support cells in fly ovaries, and while they are not germline cells, they do express a uniquely tailored version of the piRNA pathway. Out of the three Piwi proteins in *Drosophila*, only Piwi is expressed in follicular cells. In these cells it specifically loads with piRNA generated from the unidirectional cluster, *Flamenco* [20]. The *Flamenco* cluster appears to primarily target the LTR retrotransposon, *Gypsy*, whose expression can lead to the formation of viral particles that might infect the germ line [51]. Therefore, transcriptional repression of *Gypsy* loci by Piwi in follicular cells is crucial for germline survival.

In summary, while little is known about the mechanism of piRNA-mediated transcriptional silencing, it is apparent that both in mouse and in *Drosophila* germ cells, Piwi proteins play an essential role in inhibiting TE transcription and establishing a repressive genomic environment.

### ***Germ granules: components, structure and function***

Molecules involved in the same cellular process often assemble into granules—membraneless subcellular compartments—in order to increase the local concentration of

factors and enhance the efficiency and specificity of biological processes. Germ granules are conserved germ cell components present in most (if not all) sexually reproducing metazoans, and are essential in ensuring the reproductive potential of an individual organism [52-54].

Germ granules were discovered more than a century ago, as darkly staining structures that could be traced from gametes of one generation into germ cells of the next in the fly [55]. The importance of these granules for embryonic germ cell development and fertility was highlighted with experiments showing that their destruction by UV irradiation resulted in the so-called grandchild-less phenotype— sterility of the next generation [56, 57]. Furthermore, transplanting germ plasm, which contains germ granules, from a healthy donor strain into oocytes that had previously been treated with UV irradiation, could rescue infertility [58].

Detailed analysis of the constituents of germ granules reveals interesting relationships. In both mouse and fly, germ granules are frequently in close association with mitochondria and their constituent proteins have functions that are linked to the piRNA pathway [59]. Interestingly, several mitochondrial proteins including Zucchini (Zuc) and GASZ are critical piRNA pathway components [60, 61]. Furthermore, many components that localize to germ granules contain Tudor domains [62]. Tudor domains bind symmetrically di-methylated arginine (sDMA) residues, such as ones present in the N-terminal region of Piwi proteins. It is believed that Tudor domain proteins might serve as a scaffold to bring Piwi proteins together with other components of the piRNA pathway in germ granules.

In flies, germ granules can be identified in almost all stages of germ cell development. Whether in testes or ovaries, granules have various names depending on their protein constituents, localization, and stage of expression. Only recently has it been uncovered that many components of the granules operate to ensure successful transposon repression.

***Nuage are composed of many piRNA pathway components in ovarian nurse cells***

In nurse cells of the fruit fly ovary, nebulous perinuclear structures termed ‘nuage’ (after the French word for ‘cloud’) are visible in negative stained electron micrographs

[53]. The constituents of these granules include many cytoplasmic piRNA pathway components, including Aub and Ago3 (Figure 3A; Table 1). The cytoplasmic functions of the piRNA pathway, such as piRNA loading and Ping-pong, are believed to occur within these granules. Deletion or delocalization of any of the piRNA pathway components from nuage leads to impairment of the pathway, indicating the significance of these granules in proper TE silencing.

The composition and requirements of nuage granule assembly is not yet fully understood. However, some general trends for nuage assembly have been worked out. First, the components of nuage granules seem to assemble in a fashion dependent on piRNA biogenesis, since components including Aub and Ago3 that would normally localize to nuage are cytoplasmically dispersed in ovaries mutant for piRNA biogenesis factors, such as cluster associated factors Rhino and Cutoff and the mRNA export protein, UAP56 [63-65]. Second, the assembly of proteins that make up nuage granules appears to be hierarchical. For example, it is apparent from mapping pair-wise genetic interactions that some protein components only localize to nuage if other factors have already ‘built’ on the nuage complex. Some proteins may not localize to nuage in the absence of one or more upstream protein components, while the localization of these upstream components is unperturbed by mutations in downstream components [20, 41, 66, 67].

Lastly, arginine rich motifs typically found near the N-termini of Piwi proteins are modified to symmetrical di-methyl arginines (sDMA) by the PRMT5 methylosome complex [11, 13, 68, 69]. Methylated arginines can bind Tudor domains, which are a prevalent domain found in many piRNA pathway components (Table 1)[69-71]. It is believed that Tudor domain proteins might form a scaffold in nuage that helps bring together functional components of the pathway. While the role of arginine methylation in the function of the piRNA pathway is not entirely clear, Aub requires sDMA modifications to accumulate at the pole plasm during oogenesis, indicating its requirement for proper localization. Additionally, sDMA modifications on Ago3 appear to be required for interactions with another component, Papi [72]. Chapter 4 of this thesis discusses in greater detail how Piwi proteins are regulated by arginine methylation.

### ***The pole plasm: germ granules in the oocyte***

The pole plasm is defined by an accumulation of maternally deposited RNA, mitochondria and piRNA-associated granules at the posterior pole of the oocyte. Much of the material that adheres to the pole plasm will be inherited by pole cells, which are among the first cells to cellularize during embryogenesis[73]. Pole cells migrate during gastrulation to the mid-gut and give rise to primordial germ cells of the developing larva. Aub is among the hand-full of piRNA pathway proteins that localize to the pole plasm (Figure 3B; Table 1). The enrichment of Aub at the pole plasm depends on arginine methylation that allows it to bind Tudor, which itself is anchored to the pole plasm by the PRMT5 methylosome complex and the body-patterning-associated factor, Oskar [68, 74-79].

The accumulation of Aub at the pole plasm by a devoted mechanism (i.e., post-translational sDMA modification) suggests that Aub is required for the biology of primordial germ cells. Whether the role of Aub in primordial germ cells is solely palliative in destroying transposable element transcripts, as in the context of Ping-Pong, or whether its role in primordial germ cells serves a deeper purpose— one required for the establishment of the primordial gonad— remains to be understood. While an establishment of literature seems to uphold the idea that Piwi and Ago3 might also accumulate at the pole plasm, clear evidence regarding biochemical interactions of either Piwi or Ago3 with any pole plasm components is still lacking [13, 80].

### ***Suppressor of Stellate is a specialized piRNA mechanism in Drosophila Testis***

Beyond transposon control, the piRNA pathway seems to have been adopted for alternative functions: *Drosophila* testes employ a unique piRNA pathway mechanism whose primary role is to provide “quality control” during spermatogenesis. The *Stellate* locus on chromosome X is comprised of a tandemly repeated gene that encodes a protein, Stellate, whose hyper-expression is responsible for the formation of large crystalline needle-like fibers that are toxic to the cell. Due to their high copy number, unchecked expression of *Stellate* kills primary spermatocytes and results in male sterility. The antidote to *Stellate* toxicity is provided by the *Crystal* locus on chromosome Y, which encodes another set of tandem repeats named *Suppressor-of-Stellate* (*Su(Ste)*) [81, 82]. Interestingly, the *Su(Ste)* gene shares 90% homology to *Stellate* and produces both sense

and antisense transcripts. The antisense transcript plays an integral role in silencing *Stellate* mRNA and the mechanism of silencing requires many of the same components as the piRNA pathway in ovaries.

Repression of *Stellate* in testes requires Aub and Ago3, while Piwi appears to be dispensable [40, 81, 83, 84]. Additionally, the testes of mutants for piRNA pathway components SpnE, Qin, Armi, Tej, and Vret, also express *Stellate* and/or appear to have meiotic abnormalities [42, 67, 84-87]. The piRNA pathway in testes also directs TE repression through both Aub and Piwi, with Piwi being required for fertility in males [7, 84]. It is unlikely that Ping-Pong plays a significant role in testis biology, since Ago3 mutant males are semi-fertile (even though crystals of *Stellate* do accumulate within spermatocytes) and male germ stem cell maintenance is only partially defective [40, 88]. Many pathway components localize to perinuclear areas of spermatocytes forming piRNA nuage giant bodies, or piNG-bodies [88]. These components likely have a similar structure to nuage in ovary.

The implied biological function of *Stellate* and *Su(Ste)* remains ambiguous. Certainly, this system does appear to ensure that spermatocytes develop with an intact piRNA pathway, in addition to preventing Y chromosome aneuploidy during spermatogenesis. Currently, a homologous “quality assurance” mechanism is not known to exist in the female germline of the fly, or in mouse. The existence of such alternative functions of the piRNA pathway raises the question whether it has unidentified functions in the germline or in other tissues. Although expression of the core components of the piRNA pathway— Piwi proteins— seems to be restricted to the germline, it can not be excluded that specified cells (e.g., a subset of cells in the brain) do express and employ the piRNA pathway for yet uncharacterized functions beyond transposon control.

## REFERENCES

1. Slotkin, R. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome*. Nature reviews. Genetics, 2007. **8**(4): p. 272-285.
2. Siomi, M., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nature reviews. Molecular cell biology, 2011. **12**(4): p. 246-258.
3. Carmell, M., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis*. Genes & development, 2002. **16**(21): p. 2733-2742.
4. Cox, D., et al., *A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal*. Genes & development, 1998. **12**(23): p. 3715-3727.
5. Kuramochi-Miyagawa, S., et al., *Mili, a mammalian member of piwi family gene, is essential for spermatogenesis*. Development (Cambridge, England), 2004. **131**(4): p. 839-849.
6. Kuramochi-Miyagawa, S., et al., *Two mouse piwi-related genes: miwi and mili*. Mechanisms of development, 2001. **108**(1-2): p. 121-133.
7. Lin, H. and A. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development (Cambridge, England), 1997. **124**(12): p. 2463-2476.
8. Houwing, S., et al., *A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish*. Cell, 2007. **129**(1): p. 69-82.
9. Song, J.-J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science (New York, N.Y.), 2004. **305**(5689): p. 1434-1437.
10. Wang, Y., et al., *Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex*. Nature, 2008. **456**(7224): p. 921-926.
11. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nature cell biology, 2009. **11**(5): p. 652-658.
12. Liu, K., et al., *Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(43): p. 18398-18403.



13. Nishida, K., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. The EMBO journal, 2009. **28**(24): p. 3820-3831.
14. Vagin, V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes & development, 2009. **23**(15): p. 1749-1762.
15. Huang, H.-Y., et al., *Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish*. The EMBO journal, 2011. **30**(16): p. 3298-3308.
16. Mathioudakis, N., et al., *The multiple Tudor domain-containing protein TDRD1 is a molecular scaffold for mouse Piwi proteins and piRNA biogenesis factors*. RNA (New York, N.Y.), 2012. **18**(11): p. 2056-2072.
17. Aravin, A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Molecular cell, 2008. **31**(6): p. 785-799.
18. Aravin, A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS genetics, 2009. **5**(12).
19. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-1103.
20. Malone, C.D., et al., *Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary*. Cell, 2009. **137**(3): p. 522-35.
21. Aravin, A., et al., *A novel class of small RNAs bind to MILI protein in mouse testes*. Nature, 2006. **442**(7099): p. 203-207.
22. Girard, A., et al., *A germline-specific class of small RNAs binds mammalian Piwi proteins*. Nature, 2006. **442**(7099): p. 199-202.
23. Grivna, S., et al., *A novel class of small RNAs in mouse spermatogenic cells*. Genes & development, 2006. **20**(13): p. 1709-1714.
24. Lau, N., et al., *Characterization of the piRNA complex from rat testes*. Science (New York, N.Y.), 2006. **313**(5785): p. 363-367.
25. Aravin, A.A., et al., *Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control*. Science, 2007. **316**(5825): p. 744-747.
26. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes & development, 2013. **27**(4): p. 390-399.

27. Sienski, G., D. Dönertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-980.
28. Ro, S., et al., *Cloning and expression profiling of testis-expressed piRNA-like RNAs*. RNA (New York, N.Y.), 2007. **13**(10): p. 1693-1702.
29. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Molecular cell, 2013. **50**(5): p. 749-761.
30. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Molecular cell, 2013. **50**(5): p. 762-777.
31. Muerdter, F., et al., *A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila*. Molecular cell, 2013. **50**(5): p. 736-748.
32. Olivieri, D., et al., *An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila*. The EMBO journal, 2010. **29**(19): p. 3301-3317.
33. Ipsaro, J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-283.
34. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-287.
35. Voigt, F., et al., *Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc*. RNA (New York, N.Y.), 2012. **18**(12): p. 2128-2134.
36. Szakmary, A., et al., *The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster*. The Journal of Cell Biology, 2009. **185**.
37. Kawaoka, S., et al., *3' end formation of PIWI-interacting RNAs in vitro*. Molecular cell, 2011.
38. Gunawardane, L.S., et al., *A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila*. Science, 2007. **315**(5818): p. 1587-90.

39. Huang, H., et al., *AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification*. J Cell Biol, 2014. **206**(2): p. 217-30.
40. Li, C., et al., *Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies*. Cell, 2009. **137**(3): p. 509-21.
41. Anand, A. and T. Kai, *The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila*. EMBO J, 2012. **31**(4): p. 870-82.
42. Zhang, Z., et al., *Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains*. Mol Cell, 2011. **44**(4): p. 572-84.
43. Xiol, J., et al., *RNA clamping by vasa assembles a piRNA amplifier complex on transposon transcripts*. Cell, 2014. **157**(7): p. 1698-711.
44. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline*. Curr Biol, 2001. **11**(13): p. 1017-27.
45. Bozzetti, M.P., et al., *The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2*. Proc Natl Acad Sci U S A, 1995. **92**(13): p. 6067-71.
46. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-760.
47. Vagin, V.V., et al., *A distinct small RNA pathway silences selfish genetic elements in the germline*. Science, 2006. **313**(5785): p. 320-4.
48. Aravin, A.A., et al., *Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line*. Mol Cell Biol, 2004. **24**(15): p. 6742-50.
49. Handler, D., et al., *A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors*. EMBO J, 2011. **30**(19): p. 3977-93.
50. Patil, V.S. and T. Kai, *Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas*. Curr Biol, 2010. **20**(8): p. 724-30.

51. Stapleton, W., S. Das, and B.D. McKee, *A role of the Drosophila homeless gene in repression of Stellate in male meiosis*. Chromosoma, 2001. **110**(3): p. 228-40.
52. Kibanov, M.V., et al., *A novel organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing*. Mol Biol Cell, 2011. **22**(18): p. 3410-9.
53. al-Mukhtar, K.A. and A.C. Webb, *An ultrastructural study of primordial germ cells, oogonia and early oocytes in Xenopus laevis*. J Embryol Exp Morphol, 1971. **26**(2): p. 195-217.
54. Eddy, E.M. and S. Ito, *Fine structural and radioautographic observations on dense perinuclear cytoplasmic material in tadpole oocytes*. J Cell Biol, 1971. **49**(1): p. 90-108.
55. Mahowald, A.P., *Polar granules of Drosophila. II. Ultrastructural changes during early embryogenesis*. J Exp Zool, 1968. **167**(2): p. 237-61.
56. Hegner, R.W., *The History of the Germ Cells in the Paedogenetic Larva of Miastor*. Science, 1912. **36**(917): p. 124-6.
57. Hegner, R.W., *The Germ Cell Determinants in the Eggs of Chrysomelid Beetles*. Science, 1911. **33**(837): p. 71-2.
58. Hathaway, D.S. and G.G. Selman, *Certain aspects of cell lineage and morphogenesis studied in embryos of Drosophila melanogaster with an ultra-violet micro-beam*. J Embryol Exp Morphol, 1961. **9**: p. 310-25.
59. Okada, M., I.A. Kleinman, and H.A. Schneiderman, *Restoration of fertility in sterilized Drosophila eggs by transplantation of polar cytoplasm*. Dev Biol, 1974. **37**(1): p. 43-54.
60. Aravin, A. and D. Chan, *piRNAs meet mitochondria*. Developmental cell, 2011. **20**(3): p. 287-288.
61. Ipsaro, J.J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-83.
62. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Mol Cell, 2013. **50**(5): p. 749-61.
63. Ponting, C.P., *Tudor domains in proteins that interact with RNA*. Trends Biochem Sci, 1997. **22**(2): p. 51-2.

64. Klattenhoff, C., et al., *The Drosophila HP1 homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters*. Cell, 2009. **138**(6): p. 1137-49.
65. Pane, A., et al., *The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline*. EMBO J, 2011. **30**(22): p. 4601-15.
66. Zhang, F., et al., *UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery*. Cell, 2012. **151**(4): p. 871-84.
67. Lim, A.K. and T. Kai, *Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster*. Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6714-9.
68. Anne, J. and B.M. Mechler, *Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuleen*. Development, 2005. **132**(9): p. 2167-77.
69. Kirino, Y., et al., *Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization*. RNA, 2010. **16**(1): p. 70-8.
70. Liu, H., et al., *Structural basis for methylarginine-dependent recognition of Aubergine by Tudor*. Genes Dev, 2010. **24**(17): p. 1876-81.
71. Kirino, Y., et al., *Arginine methylation of vasa protein is conserved across phyla*. J Biol Chem, 2010. **285**(11): p. 8148-54.
72. Liu, L., et al., *PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition*. Development, 2011. **138**(9): p. 1863-73.
73. Hay, B., et al., *Identification of a component of Drosophila polar granules*. Development, 1988. **103**(4): p. 625-40.
74. Anne, J., *Targeting and anchoring Tudor in the pole plasm of the Drosophila oocyte*. PLoS One, 2010. **5**(12): p. e14362.
75. Arkov, A.L., et al., *The role of Tudor domains in germline development and polar granule architecture*. Development, 2006. **133**(20): p. 4053-62.
76. Anne, J., et al., *Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila*. Development, 2007. **134**(1): p. 137-46.

77. Thomson, T. and P. Lasko, *Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning*. *Genesis*, 2004. **40**(3): p. 164-70.
78. Kugler, J.M. and P. Lasko, *Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis*. *Fly* (Austin), 2009. **3**(1): p. 15-28.
79. Liang, L., W. Diehl-Jones, and P. Lasko, *Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities*. *Development*, 1994. **120**(5): p. 1201-11.
80. Megosh, H.B., et al., *The role of PIWI and the miRNA machinery in Drosophila germline determination*. *Curr Biol*, 2006. **16**(19): p. 1884-94.
81. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. *Nature*, 2009. **461**(7268): p. 1296-1299.
82. Klenov, M., et al., *Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline*. *Nucleic acids research*, 2007. **35**(16): p. 5430-5438.
83. Chambeyron, S., et al., *piRNA-mediated nuclear accumulation of retrotransposon transcripts in the Drosophila female germline*. *Proceedings of the National Academy of Sciences*, 2008. **105**(39): p. 14964-14969.
84. Shpiz, S., et al., *rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus*. *Nucleic acids research*, 2009. **37**(1): p. 268-278.
85. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. *Nucleic acids research*, 2011. **39**(20): p. 8703-8711.
86. Rozhkov, N., M. Hammell, and G. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. *Genes & development*, 2013. **27**(4): p. 400-412.
87. Soper, S., et al., *Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis*. *Developmental cell*, 2008. **15**(2): p. 285-297.
88. Song, S.U., et al., *Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the gypsy retroelement of Drosophila*. *Development*, 1997. **124**(14): p. 2789-98.

89. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(46): p. 18760-5.
90. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-9.
91. Aravin, A.A., G.J. Hannon, and J. Brennecke, *The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race*. Science, 2007. **318**(5851): p. 761-4.
92. Siomi, M.C., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nat Rev Mol Cell Biol, 2011. **12**(4): p. 246-58.
93. Ma, J.B., K. Ye, and D.J. Patel, *Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain*. Nature, 2004. **429**(6989): p. 318-22.
94. Ma, J.B., et al., *Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein*. Nature, 2005. **434**(7033): p. 666-70.
95. Yuan, Y.R., et al., *Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage*. Mol Cell, 2005. **19**(3): p. 405-19.
96. Song, J.J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science, 2004. **305**(5689): p. 1434-7.
97. Lin, H. and A.C. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development, 1997. **124**(12): p. 2463-76.
98. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-60.
99. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-103.
100. Sienski, G., D. Donertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-80.

101. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes Dev, 2013. **27**(4): p. 390-9.
102. Rozhkov, N.V., M. Hammell, and G.J. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes Dev, 2013. **27**(4): p. 400-12.
103. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic Acids Res, 2011. **39**(20): p. 8703-11.
104. Eddy, E.M., *Germ plasm and the differentiation of the germ cell line*. Int Rev Cytol, 1975. **43**: p. 229-80.
105. Mahowald, A.P., *Polar granules of drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis*. J Exp Zool, 1971. **176**(3): p. 345-52.
106. Patil, V.S. and T. Kai, *Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas*. Curr Biol, 2010.
107. Findley, S.D., et al., *Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage*. Development, 2003. **130**(5): p. 859-71.
108. Lim, A.K., L. Tao, and T. Kai, *piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline*. J Cell Biol, 2009. **186**(3): p. 333-42.
109. Nishida, K.M., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. EMBO J, 2009. **28**(24): p. 3820-31.
110. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nat Cell Biol, 2009. **11**(5): p. 652-8.
111. Schupbach, T. and E. Wieschaus, *Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology*. Genetics, 1991. **129**(4): p. 1119-36.
112. Harris, A.N. and P.M. Macdonald, *Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C*. Development, 2001. **128**(14): p. 2823-32.
113. Bossing, T., C.S. Barros, and A.H. Brand, *Rapid tissue-specific expression assay in living embryos*. Genesis, 2002. **34**(1-2): p. 123-6.



114. Djuranovic, S., et al., *Allosteric regulation of Argonaute proteins by miRNAs*. Nat Struct Mol Biol, 2010. **17**(2): p. 144-50.
115. Parker, J.S., S.M. Roe, and D. Barford, *Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex*. Nature, 2005. **434**(7033): p. 663-6.
116. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-7.
117. Snee, M.J. and P.M. Macdonald, *Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components*. J Cell Sci, 2004. **117**(Pt 10): p. 2109-20.
118. Nagao, A., et al., *Gender-Specific Hierarchy in Nuage Localization of PIWI-Interacting RNA Factors in Drosophila*. Front Genet, 2011. **2**: p. 55.
119. Wang, Y., et al., *Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes*. Nature, 2009. **461**(7265): p. 754-61.
120. Aravin, A.A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS Genet, 2009. **5**(12): p. e1000764.
121. Aravin, A.A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Mol Cell, 2008. **31**(6): p. 785-99.
122. Shoji, M., et al., *The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline*. Dev Cell, 2009. **17**(6): p. 775-87.
123. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Mol Cell, 2013. **50**(5): p. 762-77.
124. Vermaak, D., S. Henikoff, and H.S. Malik, *Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila*. PLoS Genet, 2005. **1**(1): p. 96-108.
125. Obbard, D.J., et al., *The evolution of RNAi as a defence against viruses and transposable elements*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1513): p. 99-115.
126. Simkin, A., et al., *Recurrent and recent selective sweeps in the piRNA pathway*. Evolution, 2013. **67**(4): p. 1081-90.

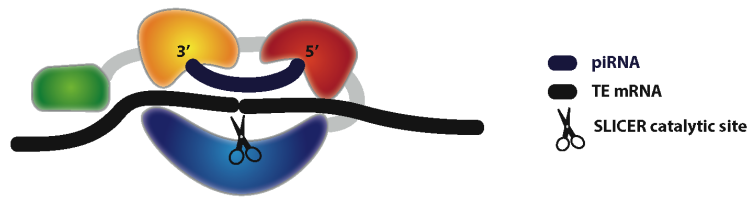
127. Comeron, J.M., *A method for estimating the numbers of synonymous and nonsynonymous substitutions per site*. J Mol Evol, 1995. **41**(6): p. 1152-9.
128. Anne, J., *C-terminal moiety of Tudor contains its in vivo activity in Drosophila*. PLoS One, 2010. **5**(12): p. e14378.
129. Creed, T.M., et al., *Novel role of specific Tudor domains in Tudor-Aubergine protein complex assembly and distribution during Drosophila oogenesis*. Biochem Biophys Res Commun, 2010. **402**(2): p. 384-9.
130. Vagin, V.V., G.J. Hannon, and A.A. Aravin, *Arginine methylation as a molecular signature of the Piwi small RNA pathway*. Cell Cycle, 2009. **8**(24): p. 4003-4.
131. Arkov, A.L. and A. Ramos, *Building RNA-protein granules: insight from the germline*. Trends Cell Biol, 2010. **20**(8): p. 482-90.
132. Siomi, M.C., T. Mannen, and H. Siomi, *How does the royal family of Tudor rule the PIWI-interacting RNA pathway?* Genes Dev, 2010. **24**(7): p. 636-46.
133. Cavey, M., et al., *Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation*. Development, 2005. **132**(3): p. 459-68.
134. Vagin, V.V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes Dev, 2009. **23**(15): p. 1749-62.
135. Bedford, M.T. and S.G. Clarke, *Protein arginine methylation in mammals: who, what, and why*. Mol Cell, 2009. **33**(1): p. 1-13.
136. Krause, C.D., et al., *Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential*. Pharmacol Ther, 2007. **113**(1): p. 50-87.
137. Wang, J., et al., *Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis*. Curr Biol, 2009. **19**(8): p. 640-4.
138. Saxe, J.P., et al., *Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline*. EMBO J, 2013. **32**(13): p. 1869-85.
139. Pandey, R.R., et al., *Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16492-7.

140. Cora, E., et al., *The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs*. RNA, 2014. **20**(6): p. 773-81.

## FIGURES AND FIGURE LEGENDS

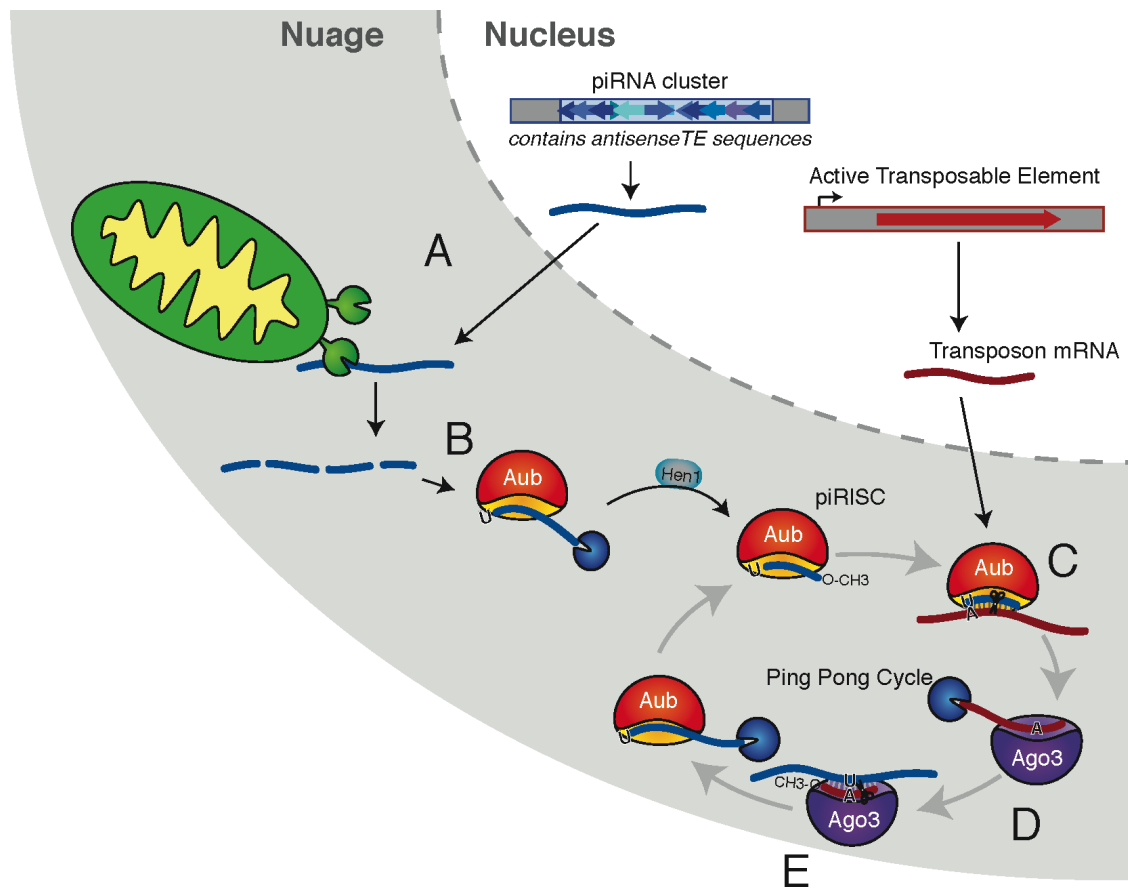
### ***Figure 1: The Domain Structure of Piwi-Clade Argonaute Proteins.***

(A) The domains of Piwi proteins are organized in a similar fashion to other Argonaute proteins, and are composed of the N-terminal region, PAZ (Piwi-Argonaute-Zwille), mid and PIWI domains. The N-terminal region consists of a notional domain that is characterized by arginine rich motifs that are targeted for methylation, and in the case of Piwi in flies and Miwi2 in mouse contains the NLS signal. (B) The organization of protein domains in space shows how the mid-domain anchors piRNA at its 5' end and the PAZ domain holds the 3' end of the piRNA. PIWI is the largest domain, and its catalytic site responsible for 'slicer' activity is positioned to cleave the backbone of annealed target RNA exactly 10 nt relative to the 5' end of the piRNA.



***Figure 2: Biogenesis of piRNA in flies begins in the nucleus and initiates the ping-ping cycle in the cytoplasm.***

(A) Precursor transcripts originating from piRNA clusters that contain anti-sense TE sequences are exported from the nucleus and processed into smaller fragments, possibly by the mitochondrial-associated endonuclease, Zucchini. (B) Aub binds precursor piRNA fragments with a preference for fragments that contain a 5' terminal uracil (1U). Mature primary piRNA is generated when an unknown 3'→5' exonuclease trims the precursor piRNA fragment bound to Aub and Hen1 catalyzes the 2-O methylation of the piRNA 3' end. (C) When mRNA of active transposons is exported from the nucleus, Aub piRNA anneals to complementary sequences within the mRNA and cleaves the transcript. (D) The cleaved transcript is loaded into Ago3, followed by trimming and 2' O-methylation of its 3' end, to generate mature secondary piRNA. (E) Ago3 loaded with the secondary piRNA is believed to target and cleave newly generated precursor transcripts that are loaded into Aub to initiate a new round of Ping-Pong.

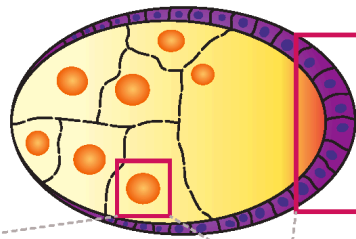


***Figure 3: Many protein factors make up Nuage granules in Drosophila nurse cells and some accumulate at the pole plasm of the developing oocyte.***

(A) Nuage granules surround the nucleus of nurse cells and consist of many factors involved in the piRNA pathway. Nuage also shows close proximity to mitochondria, which contain piRNA pathway components (Zuc, GASZ) in their outer membrane. Nuage integrity depends on nuclear piRNA biogenesis factors (such as Rhi, Cuff, UAP56) and a number of nuage components of unknown function. Many nuage components contain Tudor domains, which are receptors for symmetrically di-methylated arginines (sDMA). These Tudor proteins are believed to form a scaffold for the factors involved in the pathway. Aub and Ago3 contain sDMA motifs at their N-terminal regions and can interact with Tudor proteins. The enzyme responsible for methylation, PRMT5, consisting of Capsuléen (Csul) and Valois (Vls), is also present in nuage and associates with Tudor. Some nuclear proteins, such as Piwi and Mael, are believed to transiently visit the nuage and this is required for their function [89, 90]. Known interactions are indicated. (B) At later stages of oogenesis, nurse cells deposit their cytoplasmic contents, including piRNA pathway components, into the oocyte. Some components, including sDMA-modified Aub, accumulate at the posterior end of oocytes to form the pole plasm. The material that accumulates at the pole plasm becomes the cytoplasmic material of pole cells, which give rise to primordial germ cells of the embryo.

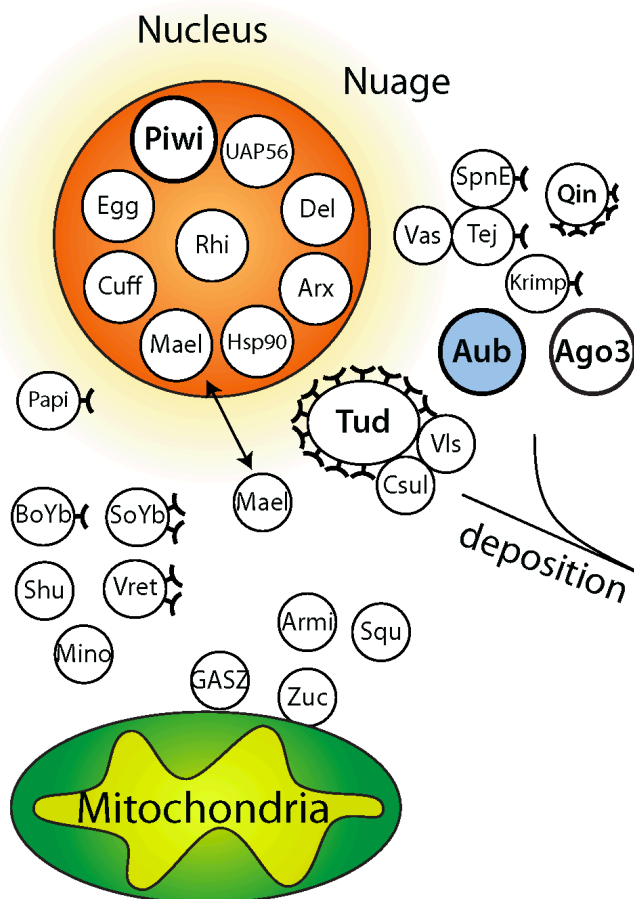


# Egg Chamber



A

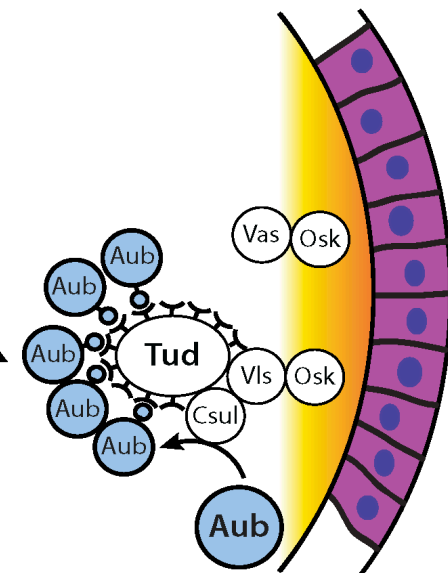
Nurse Cells



B

Oocyte

Pole Plasm



***Table 1: Proteins with identified roles in the Drosophila piRNA pathway.***

List of proteins and corresponding genes identified in the literature with functions associated with normal piRNA pathway function. Proteins are grouped into Piwi proteins, Nuclear factors, Cytoplasmic non-Tudor proteins, and Tudor proteins. Putative domains are listed for *Drosophila* and in parenthesis are domains corresponding to known homologs in mouse (M). Putative functions are based on findings presented in the referenced literature. The column listed 'Mouse' indicates the name of the homologous protein in mouse.

Piwi Proteins						
Gene	Synonyms	Tissue	Localization	Domains (M)	Function	Mouse
Piwi	<i>Pwi</i>	Germ; Soma	Nucleus	PAZ; PIWI	Argonaute; RISC effector	Ref.
Aubergine	<i>Aub</i> ; <i>Sting</i>	Germ	Cyto.: Nuage, Pole Plasm	PAZ; PIWI	Argonaute; RISC effector	Cox et al., 1998; Lin & Spradling, 1997
Argonaute-3	<i>Ago3</i>	Germ	Cyto.: Nuage	PAZ; PIWI	Argonaute; RISC effector	Schmidt et al., 1999; Harris & Macdonald, 2001 Li et al., 2009

Nuclear Factors						
Gene	Synonyms	Tissue	Localization	Domains (M)	Function	Mouse
Rhino	<i>Rhi</i> ; <i>HP1D</i>	Germ	Nucleus	Chromo; Shadow	Chromatin Factor	Ref.
Egless	<i>Egg</i> ; <i>SetDB1</i>	Germ	Nucleus	DOM3Z	H3K9me3 HMT	Volpe et al., 2001; Klattenhoff et al., 2009
Cutoff	<i>Cuff</i>	Germ	Nucleus	DOM3Z		Rangan et al., 2011
Deadlock	<i>Del</i>	Germ	Nucleus	-		Pane et al., 2011
UAP56	<i>Hei25E</i>	Germ; Soma	Nucleus	DEXDc; Hel-C	RNA Helicase	Czech et al., 2013
Asterix	<i>Arx</i> ; <i>CG3893</i>	Germ; Soma	Nucleus	CHHC; ZnF x 2	HMT cofactor	Zhang et al., 2012
Maelstrom	<i>Mael</i>	Germ	Nucleus & Cyto.	HMG		Muerdter et al., 2013
Hsp90	<i>Hsp93</i>	Germ; Soma	Nucleus & Cyto.		Heat-Shock Response; Chaperone	Finley et al., 2003; Sienski et al., 2012; Sato et al., 2011 Olivieri et al., 2012; Zhang et al., 2012

Cytoplasmic Non-Tudor Proteins						
Gene	Synonyms	Tissue	Localization	Domains (M)	Function	Mouse
Armitage	<i>Armi</i>	Germ; Soma	Cyto.	SDE3		Ref.
Shutdown	<i>Shu</i>	Germ; Soma	Cyto.: Nuage	PIPISE; TPR	Co-chaperone; piRNA loading	Cook et al., 2004
Hen-1	<i>Hen1</i> ; <i>PIH1ET</i>	Germ; Soma	Cyto.	MeT	RNA 2' O-methyl transferase	Olivieri et al., 2012; Preall et al., 2012
Squash	<i>Squ</i>	Germ	Cyto.: Nuage	RNAase III	RNA nuclease	Saito et al., 2007; Kirino & Mourelatos, 2007
Zucchini	<i>Zuc</i>	Germ; Soma	Cyto.: Mito.	HKD	RNA nuclease	Pane et al., 2007
Minotaur	<i>Minot</i> ; <i>CG5508</i>	Germ; Soma	Cyto.: Mito.	Ankyrin x 3; SAM-2	Glycerol-3PO-acyltransferase (GPAT)	Pane et al., 2007; Ipsaro et al., 2012
GASZ	<i>GASZ</i>	Germ; Soma	Cyto.: Mito.	DEXDc		Vagin et al., 2013
Vasa	<i>Vos</i>	Germ	Cyto.: Nuage, Pole Plasm	PRMT5	RNA Helicase	Ma et al., 2009; Handler et al., 2013; Czech et al., 2013
Capsule	<i>Cou</i> ; <i>dART5</i> ; <i>PRMT5</i>	Germ	Cytoplasm	MEP50; WD x 4	Arginine methyl transferase	Li et al., 1994
Validis	<i>Val</i>	Germ	Cytoplasm		PRMT5 co-factor	Anne & Mechler, 2005; Anne et al., 2007
Polo		Germ	Cyto.: Nuage		Kinase	Cavey et al., 2004 Pek et al., 2012

Tudor Proteins						
Gene	Synonyms	Tissue	Localization	Domains (M)	Function	Mouse
Tudor	<i>Tud</i>	Germ	Cyto.: Nuage, Pole Plasm	Tud x11 (Tud x 8)		Ref.
Qin	<i>Qun</i>	Germ	Cyto.: Nuage	ZnF RING E3 Ligase; Tud x 5	E3 Ligase	Arkov et al., 2006; Anne, 2010
Papi	<i>Papi</i>	Germ	Cyto.: Nuage	KH x 2; Tud x1		Annand et al., 2011; Zhang et al., 2011
Tejas	<i>Tej</i>	Germ	Cyto.: Nuage	LOTUS; Tud x1		Li et al., 2011
CG9925		Germ	Cyto.: Nuage	LOTUS; DSRM; LOTUS; Tud x 3		Padli et al., 2010
CG9925		Germ	Cyto.: Nuage	ZnF-MYND; Tud x 3		Handler et al., 2011
Spindle-E	<i>SpnE</i> ; <i>Homeless</i> ; <i>Hls</i>	Germ	Cyto.: Nuage	DEXDc; Hel-C; HA2; OB; RRM; ZnF; Tud x1	RNA Helicase	Handler et al., 2011
Krimper	<i>Krimper</i> ; <i>Montecarlo</i>	Germ	Cyto.: Nuage	Tud x1		Gillespie & Berg, 1995
Fs(1)Yb	<i>Yb</i>	Soma	Cyto.	Hel-C like; DEAD; Hel-C; ZnF; Tud x 1		Lim & Kai, 2007
Sister of Yb	<i>Sorb</i>	Germ; Soma	Cyto.: Nuage	Tud x 2; DEAD; Hel-C; ZnF; CS		Szakmany et al., 2009
Brother of Yb	<i>BoYb</i>	Germ; Soma	Cyto.: Nuage	Tud x1; DEAD; Hel-C; ZnF; CS		Handler et al., 2011
Vreteno	<i>Vret</i>	Germ; Soma	Cyto.: Nuage	ZnF-MYND; RRM; Tud x 2		Handler et al., 2011 Zamparelli et al., 2011; Handler et al., 2011

## CHAPTER 2

### **Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state**

This chapter was published in:  
Genes & Development 27, 390-399 (2013)

# Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state

Adrien Le Thomas,<sup>1,2,3</sup> Alicia K. Rogers,<sup>1,3</sup> Alexandre Webster,<sup>1,3</sup> Georgi K. Marinov,<sup>1,3</sup> Susan E. Liao,<sup>1</sup> Edward M. Perkins,<sup>1</sup> Junho K. Hur,<sup>1</sup> Alexei A. Aravin,<sup>1,4</sup> and Katalin Fejes Tóth<sup>1,4</sup>

<sup>1</sup>California Institute of Technology, Pasadena, California 91125, USA; <sup>2</sup>Université Pierre et Marie Curie, Ecole Doctorale Complexité du Vivant, 75005 Paris, France

**In the metazoan germline, piwi proteins and associated piwi-interacting RNAs (piRNAs) provide a defense system against the expression of transposable elements. In the cytoplasm, piRNA sequences guide piwi complexes to destroy complementary transposon transcripts by endonucleolytic cleavage. However, some piwi family members are nuclear, raising the possibility of alternative pathways for piRNA-mediated regulation of gene expression. We found that *Drosophila* Piwi is recruited to chromatin, colocalizing with RNA polymerase II (Pol II) on polytene chromosomes. Knockdown of Piwi in the germline increases expression of transposable elements that are targeted by piRNAs, whereas protein-coding genes remain largely unaffected. Derepression of transposons upon Piwi depletion correlates with increased occupancy of Pol II on their promoters. Expression of piRNAs that target a reporter construct results in a decrease in Pol II occupancy and an increase in repressive H3K9me3 marks and heterochromatin protein 1 (HP1) on the reporter locus. Our results indicate that Piwi identifies targets complementary to the associated piRNA and induces transcriptional repression by establishing a repressive chromatin state when correct targets are found.**

[**Keywords:** piRNA; Piwi; chromatin; RNA polymerase II; transcription; transposon]

Supplemental material is available for this article.

Received November 8, 2012; revised version accepted January 14, 2013.

Diverse small RNA pathways function in all kingdoms of life, from bacteria to higher eukaryotes. In eukaryotes, several classes of small RNA associate with members of the Argonaute protein family, forming effector complexes in which the RNA provides target recognition by sequence complementarity, and the Argonaute provides the repressive function. Argonaute–small RNA complexes have been shown to regulate gene expression both transcriptionally and post-transcriptionally. Post-transcriptional repression involves cleavage of target RNA through either the endonucleolytic activity of Argonautes or sequestering targets into cytoplasmic ribonucleoprotein (RNP) granules (Hutvagner and Simard 2008).

The mechanism of transcriptional repression by small RNAs has been extensively studied in fission yeast and plants. Several studies showed that Argonaute–small RNA complexes induce transcriptional repression by tethering chromatin modifiers to target loci. In fission yeast,

the effector complex containing the Argonaute and the bound siRNA associates with the histone H3 Lys 9 (H3K9) methyltransferase Ctr4 to install repressive H3K9-dimethyl marks at target sites (Nakayama et al. 2001; Maison and Almouzni 2004; Sugiyama et al. 2005; Grewal and Jia 2007). Methylation of histone H3K9 leads to recruitment of the heterochromatin protein 1 (HP1) homolog Swi6, enhancing silencing and further promoting interaction with the Argonaute complex. The initial association of Ago with chromatin, however, requires active transcription (Ameyar-Zazoua et al. 2012; Keller et al. 2012). Plants also use siRNAs to establish repressive chromatin at repetitive regions. Contrary to yeast, heterochromatin in plants is marked by DNA methylation, although repression also depends on histone methylation by a Ctr4 homolog (Soppe et al. 2002; Onodera et al. 2005). Although siRNA-mediated gene silencing is predominant on repetitive sequences, it is not limited to these sites. Constitutive expression of dsRNA mapping to promoter regions results in production of corresponding siRNAs, de novo DNA methylation, and gene silencing (Mette et al. 2000; Matzke et al. 2004).

In metazoans, small RNA pathways are predominantly associated with post-transcriptional silencing. One class

<sup>3</sup>These authors contributed equally to this work.

<sup>4</sup>Corresponding authors  
E-mail kft@caltech.edu  
E-mail aravin@caltech.edu

Article published online ahead of print. Article and publication date are online at <http://www.genesdev.org/cgi/doi/10.1101/gad.209841.112>.

of small RNA, microRNA, regulates expression of a large fraction of protein-coding genes (Friedman et al. 2009). In *Drosophila*, siRNAs silence expression of transposable elements (TEs) in somatic cells (Chung et al. 2008; Ghildiyal et al. 2008) and target viral genes upon infection (Galiana-Arnoux et al. 2006; Wang et al. 2006; Zamboni et al. 2006). Another class of small RNAs, Piwi-interacting RNAs (piRNAs), associates with the Piwi clade of Argonautes and acts to repress mobile genetic elements in the germline of both *Drosophila* and mammals (Siomi et al. 2011). Analysis of piRNA sequences in *Drosophila* revealed a very diverse population of small RNAs that primarily maps to transposon sequences and is derived from a number of heterochromatic loci called piRNA clusters, which serve as master regulators of transposon repression (Brennecke et al. 2007). Additionally, a small fraction of piRNAs seems to be processed from the mRNA of several host protein-coding genes (Robine et al. 2009; Saito et al. 2009). The *Drosophila* genome encodes three piwi proteins: Piwi, Aubergine (AUB), and Argonaute3 (AGO3). In the cytoplasm, AUB and AGO3 work together to repress transposons through cleavage of transposon transcripts, which are recognized through sequence complementarity by the associated piRNAs (Vagin et al. 2006; Agger et al. 2007; Brennecke et al. 2007; Gunawardane et al. 2007).

In both *Drosophila* and mammals, one member of the Piwi clade proteins localizes to the nucleus. Analogously to small RNA pathways in plants, the mouse piRNA pathway is required for de novo DNA methylation and silencing of TEs (Carmell et al. 2007; Aravin et al. 2008; Kuramochi-Miyagawa et al. 2008); however, the exact mechanism of this process is unknown. In *Drosophila*, DNA methylation is absent; however, several studies indicate that elimination of Piwi from the nucleus causes changes in histone marks on TEs (Klenov et al. 2011; Pöyhönen et al. 2012), yet a genome-wide analysis of Piwi's effect on chromatin marks and transcription is lacking.

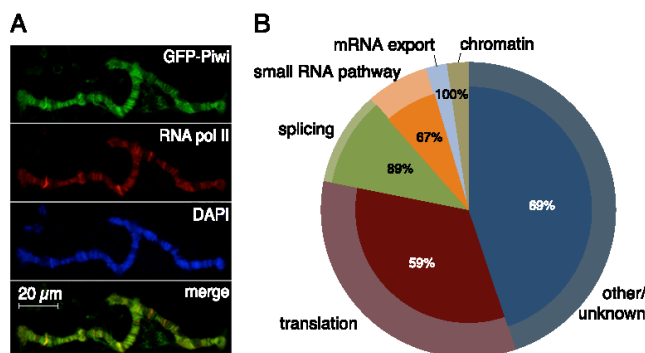
Here we show that Piwi interacts with chromatin on polytene chromosomes in nurse cell nuclei. We found that Piwi exclusively represses loci that are targeted by piRNAs. We show that Piwi-mediated silencing occurs

through repression of transcription and correlates with installment of repressive chromatin marks at targeted loci.

## Results

To analyze the role of Piwi in the nucleus, we generated transgenic flies expressing a GFP-tagged Piwi protein (GFP-Piwi) under the control of its native regulatory region. GFP-Piwi was expressed in the ovary and testis in a pattern indistinguishable from the localization of native Piwi and was able to rescue the piwi-null phenotype as indicated by ovarian morphology, fertility, transposon expression, and piRNA levels. GFP-Piwi was deposited into the mature egg and localized to the pole plasm; however, contrary to a previous observation (Brower-Toland et al. 2007), we did not detect Piwi expression outside of the ovary and testis in third instar larvae or adult flies. We also did not observe the association of Piwi with polytene chromosomes in salivary gland cells of third instar larvae. In both follicular and germline cells of the *Drosophila* ovary, GFP-Piwi localized exclusively in the nucleus, with slightly higher concentrations apparent in regions enriched for DAPI, indicating a possible interaction with chromatin. To gain further insight into Piwi localization in the nucleus, we took advantage of the fact that nurse cell chromosomes are polytenized and can be visualized on the *otu* mutant background (Mal'ceva et al. 1997). Analysis of polytene chromosomes from nurse cells demonstrated that GFP-Piwi associates with chromatin in a specific banding pattern. Interestingly, coimmunostaining showed that a GFP-Piwi signal on polytene chromosomes generally overlaps with the RNA polymerase II (Pol II) signal, which marks sites of active transcription (Fig. 1A).

In order to identify factors that might be responsible for targeting Piwi to chromatin, we immunoprecipitated Piwi complexes from the *Drosophila* ovary and analyzed Piwi interaction partners by mass spectrometry. We purified Piwi complexes from ovaries of three different transgenic lines expressing GFP-Piwi, myc-Piwi, or Flag-Piwi using antibodies against each respective tag. As a control, we used flies expressing free GFP in the ovary.

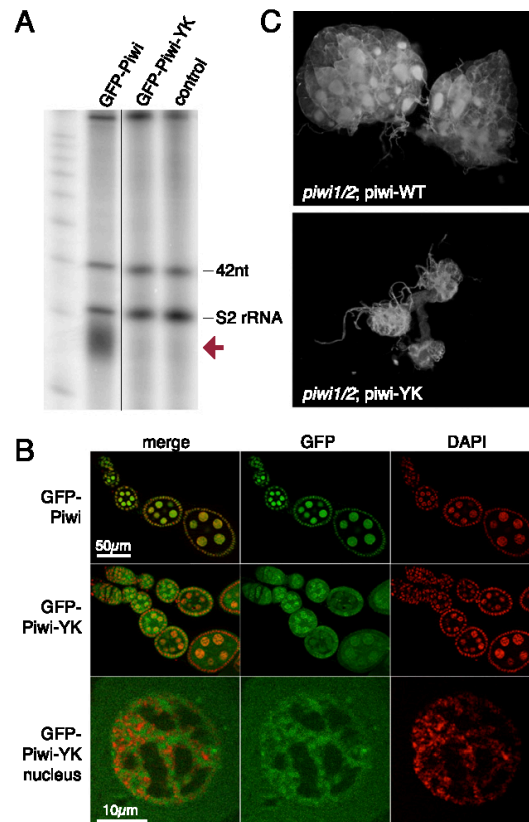


**Figure 1.** Piwi associates with chromatin and nuclear transcripts. **(A)** Polytene chromosomes from *Drosophila* nurse cells expressing GFP-Piwi on the *otu*[7]/*otu*[11] background. Piwi pattern on chromosomes correlates with Pol II staining. **(B)** Mass spectrometry analysis of Piwi interaction partners. Piwi complexes were precipitated in the presence and absence of RNase A. The outer circle represents classification of Piwi-associated proteins based on GO term analysis. The inner pies represent the fraction of each group whose association with Piwi depends on RNA (percentage indicated). Note that chromatin, splice, and mRNA export factors are virtually absent after RNase A treatment.

We identified >50 factors that showed significant enrichment in all three Piwi purifications but were absent in the control. We were unable to identify chromatin-associated factors that directly associate with Piwi but identified several RNA-binding proteins that associate with nascent transcripts, such as splicing (Rm62, Pep, Ref1, Yps, CG9684, CG31368, CG5728, and Mago) and nuclear export (Tho2 and Hpr1) factors (Fig. 1B). Upon RNase A treatment prior to immunoprecipitation, the presence of most of these RNA-binding proteins in purified Piwi complexes was eliminated.

Piwi proteins are believed to find their targets through sequence complementarity of the associated piRNA. In fact, it has been proposed that lack of the associated piRNA leads to destabilization of piwi proteins and to Piwi's inability to localize to the nucleus (Saito et al. 2009; Haase et al. 2010; Olivieri et al. 2010; Handler et al. 2011; Ishizu et al. 2011). On the other hand, Piwi has been proposed to have functions that are independent of its role in transposon control by regulating stem cell niche development (Cox et al. 1998; Klenov et al. 2011). To address the role of piRNA in translocation of Piwi into the nucleus and its function, we generated transgenic flies expressing a point mutant Piwi—referenced as Piwi-YK—that is deficient in piRNA binding due to a substitution of two conserved amino acid residues (Y551L and K555E) in the 5' phosphate-binding pocket (Kiriakidou et al. 2007; Djuranovic et al. 2010). The Piwi-YK mutant was expressed in *Drosophila* follicular and germ cells at levels similar to that of wild-type Piwi but was completely devoid of associated piRNA (Fig. 2A). In contrast to wild-type Piwi, Piwi-YK could be found in the cytoplasm, supporting the existence of a quality control mechanism that prevents entrance of unloaded Piwi into the nucleus (Ishizu et al. 2011). Nevertheless, a significant amount of piRNA-deficient Piwi localized to the nucleus (Fig. 2B). Similar to wild-type Piwi, Piwi-YK seemed to associate with chromatin, as indicated by its localization in DAPI-stained regions of the nuclei, and this is consistent with fluorescence loss in photobleaching (FLIP) experiments that demonstrated reduced nuclear mobility compared with free diffusion (Supplemental Fig. S1). Based on sterility and ovarian morphology, the *piwi-YK* transgene was unable to rescue the *piwi*-null phenotype despite its nuclear localization (Fig. 2C), indicating that while piRNA binding is not absolutely essential for stability and nuclear localization of Piwi, it is required for Piwi function.

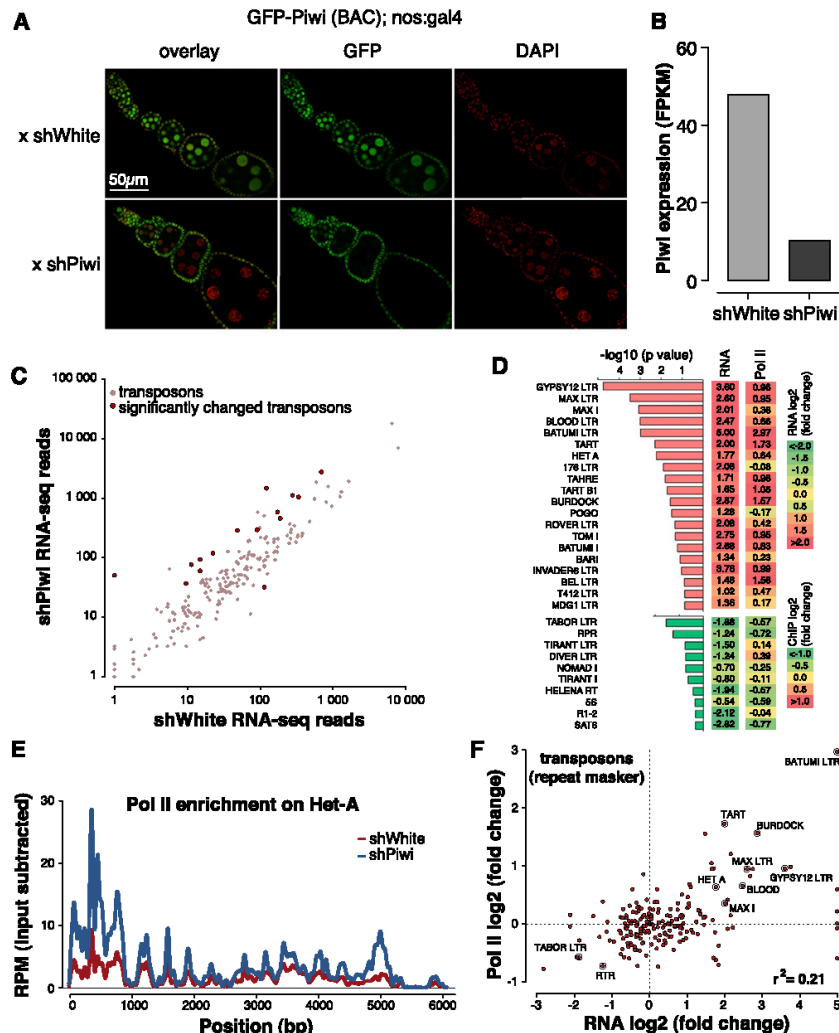
To directly test the function of Piwi in the nucleus, we analyzed the effect of Piwi deficiency on gene expression and chromatin state on a genome-wide scale. Piwi mutant females have atrophic ovaries caused by Piwi deficiency in somatic follicular cells (Lin and Spradling 1997; Cox et al. 1998), which precludes analysis of Piwi function in null mutants. Instead, we used RNAi knockdown to deplete Piwi in germ cells while leaving it functionally intact in somatic follicular cells. The Piwi knockdown flies did not exhibit gross morphological defects in the ovary; however, they showed drastic reduction in GFP-Piwi expression in germ cells and were sterile (Fig. 3A,B).



**Figure 2.** Piwi function, but not its nuclear localization, requires piRNA association. (A) The Piwi-YK mutant does not associate with piRNA. Immunoprecipitation of Piwi-piRNA complexes was performed with GFP antibody on ovaries from GFP-Piwi and GFP-Piwi-YK transgenic flies and a control strain. Small RNAs were isolated, 5'-labeled, and resolved on a denaturing gel. The same amount of 42-nucleotide RNA oligonucleotides was spiked into all samples prior to RNA isolation to control for loss of RNA during isolation and labeling. piRNAs (red arrow) are absent in the Piwi-YK complex. (B) GFP-Piwi-YK is present in the nuclei of nurse cells and colocalizes with chromatin (DAPI-stained areas). (C) The Piwi-YK mutant does not rescue the morphological changes caused by the *piwi*-null mutation. Dark-field images of ovaries where either the wild-type *piwi* or the *piwi-YK* transgene has been backcrossed onto the *piwi*-null background.

To analyze the effect of Piwi deficiency on the steady-state transcriptome as well as the transcription machinery, we performed RNA sequencing (RNA-seq) and Pol II chromatin immunoprecipitation (ChIP) combined with deep sequencing (ChIP-seq) experiments from Piwi knockdown and control flies.

In agreement with previous observations that implicated Piwi in transposon repression (Saito et al. 2006; Aravin et al. 2007; Brennecke et al. 2007), we found that steady-state transcript levels of several TEs were increased



**Figure 3.** Piwi transcriptionally represses TEs. **[A]** Piwi knockdown is efficient and specific to ovarian germ cells as indicated by GFP-Piwi localization. GFP-Piwi; Nanos-Gal4-VP16 flies were crossed to control shRNA (shWhite) or shPiwi lines. Piwi is specifically depleted in germ cells and not in follicular cells, consistent with expression of the Nanos-Gal4-VP16 driver. **[B]** Piwi expression as measured by RNA-seq in the Piwi knockdown and control lines. Note that Piwi expression is unaffected in follicular cells, leading to relatively weak apparent knockdown in RNA-seq libraries from whole ovaries. **[C]** Effect of Piwi knockdown on the expression of TEs. Two biological replicate RNA-seq experiments were carried out, and differential expression was assessed using DESeq. Transposons that show significant change ( $P < 0.05$ ) are indicated by dark-red circles. Out of 217 individual RepeatMasker-annotated TEs, 15 show a significant increase in expression upon Piwi knockdown. **[D]** The change in the levels of TE transcripts and Pol II occupancy on their promoters upon Piwi knockdown. Twenty up-regulated and 10 down-regulated transposons with the most significant changes in expression level are shown. Note the low statistical significance for down-regulated transposons. For a complete list of transposons, see Supplemental Figure S2. **[E]** Pol II signal over the Het-A retrotransposon in control flies (shWhite; red) and upon Piwi knockdown (shPiwi; blue). **[F]** Increased abundance of transposon transcripts upon Piwi depletion correlates with increased Pol II occupancy over their promoters ( $r^2 = 0.21$ ). Note that the majority of elements do not show significant change in either RNA abundance or Pol II occupancy.

upon Piwi knockdown in germ cells (Fig. 3C,D; Supplemental Fig. S2). We found little to no change of RNA levels for transposons whose activity is restricted to follicular cells of the ovary, indicating that the observed

changes are indeed due to loss of Piwi in the germline (Supplemental Fig. S2). The analysis of Pol II ChIP-seq showed that Pol II occupancy increased over promoters of multiple TEs (Fig. 3D–F; Supplemental Fig. S3). Indeed,



the change in steady-state levels of transposon transcripts upon Piwi depletion correlated with changes of Pol II occupancy (Fig. 3F). This result demonstrates that Piwi ensures low levels of transposon transcripts through a repressive effect on the transcription machinery.

To test whether Piwi-mediated transcriptional repression is accompanied by a corresponding change in chromatin state, we used ChIP-seq to analyze the genome-wide distribution of the repressive H3K9me3 mark in the ovary upon Piwi knockdown. We identified 705 genomic loci at which the level of H3K9me3 significantly decreased. More than 90% of the regions that show a decrease in the H3K9me3 mark upon Piwi depletion overlapped TE sequences, compared with the 33% that is expected from random genome sampling (Fig. 4A). Furthermore, these regions tend to be located in the heterochromatic portions of the genome that are not assembled on the main chromosomes (Fig. 4B). Only 20 of the identified regions localized to the euchromatic parts of the genome. Of these, 15 (75%) contained potentially active annotated copies of transposons. Taken together, our results indicate that Piwi is required for installment of repressive H3K9me3 chromatin marks on TE sequences of the genome.

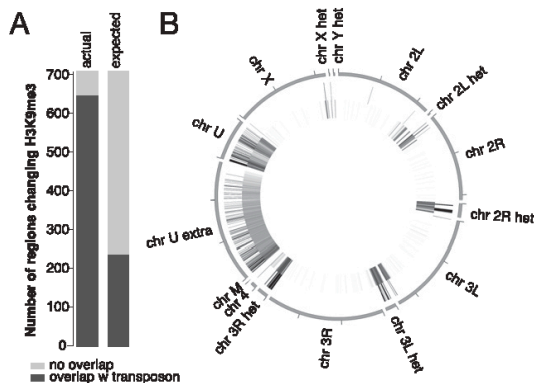
While the vast majority of protein-coding host genes did not show significant changes in transcript level or Pol II occupancy upon Piwi knockdown, the expression of a small set of protein-coding genes (150 genes with a

$P$ -value  $< 0.05$ ) was significantly increased (Fig. 5A; Supplemental Table 1). There are several possible explanations for Piwi's effect on host gene expression. First, failure in the piRNA pathway might cause up-regulation of several genes that generate piRNAs in wild-type ovaries (Robine et al. 2009; Saito et al. 2009). However, the genes up-regulated in Piwi-deficient ovaries were not enriched in piRNAs compared with other genes. Second, H3K9me3 marks installed on TE sequences in a Piwi-dependent manner might spread into neighboring host genes and repress their transcription, as was recently demonstrated in a follicular cell culture model (Sienski et al. 2012). To address this possibility, we analyzed genomic positions of the genes whose expression was increased upon Piwi knockdown relative to genomic regions that showed a decrease in H3K9me3 marks. We found that up-regulated genes did not show a significant change in the H3K9me3 mark (Fig. 5B; Supplemental Fig. S4). Furthermore, the few genes located close to the regions that show a decrease in H3K9me3 signal had unaltered expression levels upon Piwi knockdown. Next, we analyzed the functions of up-regulated genes using gene ontology (GO) term classifications and found significant enrichment for proteins involved in protein turnover and stress and DNA damage response pathways (Fig. 5C). Particularly, we found that 31 subunits of the proteasome complex were overexpressed. Therefore, our analysis indicates that up-regulation of specific host genes is likely a secondary response to elevated transposon levels and genomic damage.

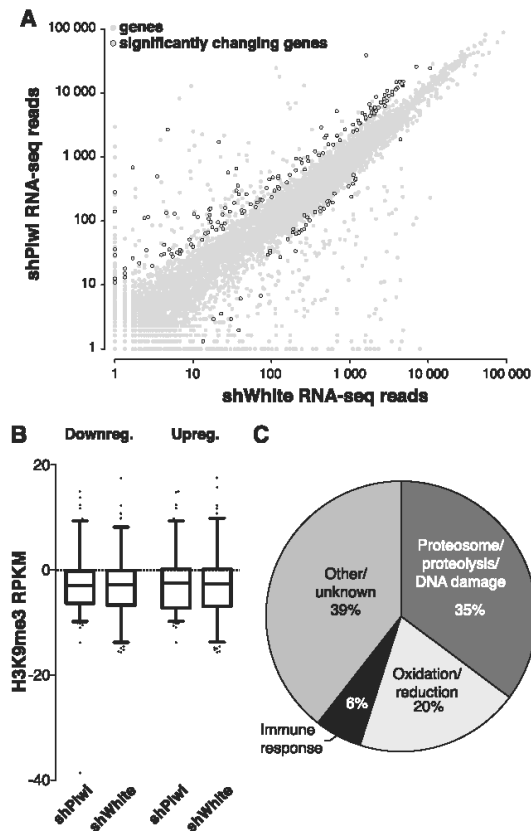
In contrast to host genes, transcripts of TEs are targeted by piRNA. To directly address the role of piRNA in Piwi-mediated transcriptional silencing, we took advantage of a fly strain that expresses artificial piRNAs against the *lacZ* gene, which are loaded into Piwi complexes and are able to repress *lacZ* reporter expression in germ cells (Fig. 6A; Josse et al. 2007; Muerdter et al. 2012). Expression of piRNAs that are antisense to the reporter gene caused transcriptional silencing of the *lacZ* gene as measured by Pol II occupancy (Fig. 6B). Furthermore, we found that piRNA-induced silencing of the reporter gene was associated with an increase in the repressive H3K9me3 mark and HP1 occupancy and a decrease in the abundance of the active H3K4me2/3 marks at the reporter locus (Fig. 6C). This result is in good agreement with the genome-wide effect of Piwi depletion on distribution of the H3K9me3 mark and suggests that transcriptional silencing correlates with the establishment of a repressive chromatin structure and is mediated by piRNAs that match the target locus.

## Discussion

Little is known about the function of nuclear piwi proteins. The nuclear piwi in mice (Miw2) affects DNA methylation of TEs (Carmell et al. 2007; Aravin et al. 2008; Kuramochi-Miyagawa et al. 2008). Several recent reports implicate *Drosophila* Piwi in regulation of chromatin marks on transposon sequences (Lin and Yin 2008; Klenov et al. 2011; Wang and Elgin 2011; Sienski et al. 2012). The mechanism of these processes is unknown in



**Figure 4.** Piwi-induced transcriptional repression correlates with establishment of a repressive chromatin state. (A) Overlap between genomic regions of H3K9me3 depletion upon Piwi knockdown and TEs. Two replicates of H3K9me3 ChIP-seq experiments were carried out on control and Piwi-depleted ovaries, and enriched regions were identified using DESeq (see the Materials and Methods for details). A total of 705 regions show significant ( $P < 0.05$ ) decrease in H3K9me3 occupancy upon Piwi knockdown, while only 30 regions showed a similarly significant increase. Out of the 705 regions that show a decrease in H3K9me3 marks upon Piwi knockdown, 91% (646) overlap with TE sequences compared with the 33% expected from random genome sampling. (B) Genomic positions of H3K9me3-depleted regions upon Piwi depletion (outer circle) and RepeatMasker-annotated transposons (inner circle). Note that almost all regions are localized in heterochromatic and repeat-rich portions of the genome (Het, chrU, and chrUExtra chromosomes).

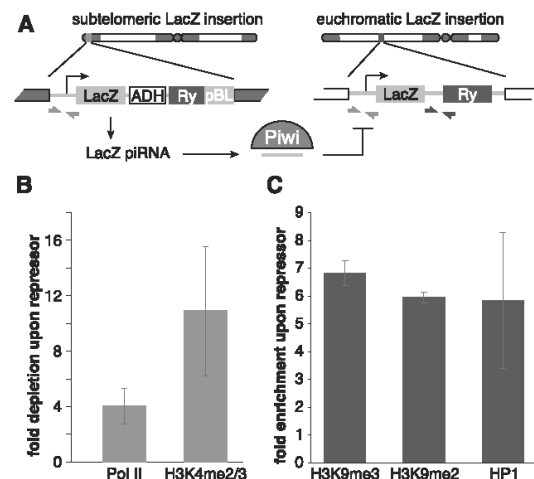


**Figure 5.** Piwi does not directly repress protein-coding genes. (A) Effect of Piwi knockdown on the expression of genes. Two replicate RNA-seq experiments were carried out, and differential expression was assessed using DESeq. Genes that show significant change ( $P < 0.05$ ) are indicated by black circles. The vast majority of genes does not change significantly upon germline Piwi knockdown (shPiwi) compared with control (shWhite). (B) H3K9me3 mark density does not change over genes that show a significant change in expression upon Piwi knockdown (see Fig. 3C). Up-regulated and down-regulated genes are plotted separately. Signal indicated is after background subtraction. (C) Functional analysis of up-regulated genes by the Database for Annotation, Visualization, and Integrated Discovery (DAVID) reveals activation of the protein degradation and DNA damage response pathways. Percentages of all up-regulated genes are indicated.

both organisms. Previously, Piwi was shown to associate with polytene chromosomes in salivary gland cells and colocalize with HP1, a chromodomain protein that binds to heterochromatin and a few loci in euchromatin, suggesting that HP1 mediates Piwi's interaction with chromatin (Brower-Toland et al. 2007). However, recent results showed that the putative HP1-binding site on Piwi is dispensable for Piwi-mediated transposon silencing (Wang and Elgin 2011).

We did not detect Piwi expression outside of the ovary and testis, including in salivary gland cells, using a GFP-

Piwi transgene expressed under native regulatory elements. We detected GFP-Piwi on polytene chromosomes in ovarian nurse cells that have a germline origin; however, it localizes in a pattern that largely does not overlap with HP1. FLIP experiments with GFP-Piwi indicated a relatively fast rate of fluorescence redistribution as compared with histone H2A (Supplemental Fig. S1), implying a transient interaction of Piwi with chromatin. Our proteomic analysis of Piwi complexes isolated from *Drosophila* ovaries did not identify chromatin-associated factors but revealed several RNA-binding proteins, such as splicing and nuclear export factors that bind nascent RNA transcripts (Fig. 1B). Importantly, the interaction of most of these RNA-binding proteins with Piwi was dependent on RNA, indicating that Piwi associates with nascent transcripts. As Piwi itself lacks DNA- and RNA-binding domains (beyond the piRNA-binding domain),



**Figure 6.** piRNA-dependent targeting of Piwi to a reporter locus leads to establishment of a repressive chromatin state and transcriptional silencing. (A) The mechanism of trans-silencing mediated by artificial piRNA and a schematic representation of the repressor and reporter *lacZ* constructs. The repressor construct is inserted in a subteleric piRNA cluster, leading to generation of piRNA from its sequence. Primers mapping to both constructs used for the Pol II and H3K4me2/3 ChIP-quantitative PCR (qPCR) are shown by light-gray arrows; primers specific to the reporter locus used for the H3K9me3, H3K9me2, and HP1 ChIP-qPCR are indicated by dark-gray arrows. (B) piRNAs induce transcriptional repression of the *lacZ* reporter. Pol II and H3K4me2/3 signals decreased on the *lacZ* promoter in the presence of artificial piRNAs as measured by ChIP-qPCR. Shown is the fold depletion of signal in flies that carry both repressor and reporter constructs compared with control flies that have only the reporter construct. The signal was normalized to RP49. (C) piRNAs induce an increase in H3K9me3 and H3K9me2 marks and HP1 binding as measured by ChIP-qPCR. Shown is the fold increase of corresponding ChIP signals downstream from the *lacZ* reporter in flies that carry both repressor and reporter constructs compared with control flies that have only reporter construct. The signal was normalized to RP49.

it is likely that the recruitment of Piwi to chromatin is through interactions with other RNA-binding proteins or sequence-specific interactions between Piwi-bound piRNA and nascent transcripts.

Using specific Piwi knockdown in germ cells of the *Drosophila* ovary, we analyzed the effect of Piwi depletion on gene expression, the transcription machinery, and H3K9me3 chromatin marks genome-wide. In agreement with previous results [Klenov et al. 2011], we found up-regulation of several TEs upon Piwi knockdown (Fig. 3C). The TEs that did not change their expression upon germline knockdown of Piwi might be expressed exclusively in somatic follicular cells of the ovary, such as the *gypsy* retrotransposon. Alternatively, some elements present in the genome might not have transcriptionally active copies, or the cytoplasmic AUB/AGO3 proteins may efficiently silence them at the post-transcriptional level.

The increase in steady-state levels of RNA upon Piwi depletion strongly correlates with an increase in Pol II occupancy on the promoters of transposons (Fig. 3D,F; Supplemental Fig S2). This result suggests that Piwi represses transposon expression at the transcriptional level, although we cannot completely exclude the possibility of an additional post-transcriptional effect. It was shown previously that depletion or mutation of Piwi leads to depletion of the repressive H3K9me3 mark and an increase in the active H3K4me2/3 marks on several transposon sequences [Klenov et al. 2011; Wang and Elgin 2011]. Our ChIP-seq data extend these results to a genome-wide scale, proving that transposons are indeed the sole targets of Piwi, and demonstrate that changes in histone marks directly correlate with transcriptional repression.

Piwi depletion in the germline does not affect expression of the majority of host genes, although a small fraction of genes changes expression (Fig. 5A). One possible mechanism of the effect Piwi has on host genes is the spreading of repressive chromatin structure from transposon sequences to adjacent host genes. Indeed, such a spreading and the resulting repression of host gene transcription were observed in an ovarian somatic cell (OSC) culture model [Sienski et al. 2012]. However, we did not find significant changes in the H3K9me3 mark for genes that are up-regulated upon germline depletion of Piwi, arguing against this mechanism playing a major role in host gene regulation. Instead, we found that the majority of host genes whose expression is increased as a result of Piwi depletion participate in protein turnover (e.g., proteasome subunits) and stress and DNA damage response pathways, indicating that they might be activated as a secondary response to cellular damage induced by transposon activation. The different effect of Piwi depletion on host gene expression in ovary and cultured cells might be explained by the fact that silencing of host genes due to transposon insertion would likely have a strong negative effect on the fitness of the organism but could be tolerated in cultured cells. Accordingly, new transposon insertions that cause repression of adjacent host genes should be eliminated from the fly population but can be detected

in cultured cells. In agreement with this explanation, the majority of cases of repressive chromatin spreading in OSCs were observed for new transposon insertions that are absent in the sequenced *Drosophila* genome. Indeed, it was shown that the vast majority of new transposon insertions is present at a low frequency in the *Drosophila* population, likely due to strong negative selection [Petrov et al. 2003]. Such selection was primarily attributed to the ability of TE sequences to cause recombination and genomic rearrangements. We propose that in addition to the effects on recombination, the selection against transposons can be driven by their negative impact on host gene expression in the germline linked to Piwi-mediated chromatin silencing.

How does Piwi discriminate its proper targets—transposons—from host genes? In the case of cytoplasmic Piwi proteins AUB and AGO3, recognition and post-transcriptional destruction of TE transcripts is guided by associated piRNAs. Our results indicate that piRNAs provide guidance for transcriptional silencing by the nuclear Piwi protein as well. First, in contrast to host genes that are not targeted by piRNAs, TE transcripts, which are regulated by Piwi, are recognized by antisense Piwi-bound piRNA [Brennecke et al. 2007]. Second, a Piwi mutant that is unable to bind piRNA failed to rescue the piwi-null mutation despite its ability to enter the nucleus. Finally, expression of artificial piRNAs that target a reporter locus induced transcriptional silencing associated with an increase in repressive H3K9me3 and HP1 chromatin marks and a decrease in the active H3K4me2/3 marks (Fig. 6B,C). In contrast, the tethering of Piwi to chromatin in a piRNA-independent fashion by fusing Piwi with the lacI DNA-binding domain that recognizes lacO sequences inserted upstream of a reporter gene did not lead to silencing of the reporter (data not shown). Together, our results demonstrate that piRNAs are the essential guides of Piwi to recognize its targets for transcriptional repression.

It is tempting to propose that, similar to Argonautes in fission yeast, *Drosophila* Piwi directly recruits the enzymatic machinery that establishes the repressive H3K9me3 mark on its targets. Establishment of repressive marks can lead to stable chromatin-based transcriptional silencing that does not require further association of Piwi with target loci. This model explains why we found that Piwi is relatively mobile in the nucleus, indicative of only a transient interaction with chromatin. The Piwi-mediated transcriptional silencing has an interesting parallel in *Caenorhabditis elegans*, where the Piwi protein PRG-1 and associated 21U RNAs are able to induce stable transgenerational repression that correlates with formation of silencing chromatin marks on target loci. Interestingly, PRG-1 and 21U RNAs are necessary only for initial establishment of silencing, while continuing repression depends on siRNA and the WAGO group of Argonautes [Ashe et al. 2012; Bagijn et al. 2012; Buckley et al. 2012; Shirayama et al. 2012]. Future studies should reveal the pathway that leads to transcriptional repression downstream from Piwi in *Drosophila* and the differences from and similarities to other species.

## Materials and methods

### *Drosophila stocks*

*Nanos-Gal4-VP16* (BL4937), *UASp-shWhite* (BL33623), *UASp-shPiwi* (BL 33724), and Chr. I and II Balancer (BL197) were purchased from the Bloomington Stock Center. GFP-Piwi-expressing flies (see below) were backcrossed onto the *piwi1/piwi2* (available from Bloomington Stock Center) background or the *otu7/otu11* (available from Bloomington Stock Center) background, respectively. *LacZ* reporter lines were a generous gift from S. Ronsseay.

### *Generation of transgenic fly lines*

The GFP-Piwi, 3xFlag-HA-Piwi, and myc-Piwi constructs were generated using bacterial recombineering (Gene Bridges Counter Selection kit) to insert the respective tag after the start codon of the Piwi genomic region cloned in BAC clone BACN04M10. The KpnI-XbaI genomic fragment that contains the Piwi gene and flanking sequences was transferred to corresponding sites of the pCasper4 vector to create pCasper4/tagged Piwi.

The pCasper4/GFP-Piwi construct was used to generate pCasper4/GFP-Piwi-YK with two point mutations, Y551I and K555E. Mutations were introduced by PCR, amplifying products corresponding to a 3.1-kb upstream fragment and a 2.58-kb downstream fragment. The upstream fragment included a unique XbaI site at the 5' end of the amplicon and overlapped 39 base pairs (bp) with the downstream fragment, which included a unique BamHI site at its 3' end. The single XbaI-BamHI fragment was generated by overlap PCR with outside primers and cloned into corresponding sites of pCasper4/GFP-Piwi to replace the wild-type fragment. Transgenic flies were generated by P-element-mediated transformation (BestGene).

### *Immunoprecipitation of Piwi proteins and RNA gel of piRNA*

Dissected ovaries were lysed in lysis buffer (20 mM HEPES at pH 7.0, 150 mM KCl, 2.5 mM MgCl<sub>2</sub>, 0.5% Triton X-100, 0.5% Igepal, 100 U/mL RNasin [Promega], EDTA-free Complete Protease Inhibitor Cocktail [Roche]) and supernatant clarified by centrifugation. Supernatant was incubated with anti-eGFP polyclonal antibody (Covance) conjugated to Protein-G Dynabeads at 4°C. Beads were spiked with 5 pmol of synthesized 42-nucleotide RNA oligomer to assess purification efficiency, proteinase K-digested, and phenol-extracted. Isolated RNA was CIP-treated, radiolabeled using PNK and  $\gamma$ -P32-labeled ATP, and run on a 15% urea-PAGE gel. Western blots of ovary lysate and anti-eGFP immunoprecipitates were obtained from 8% SDS-PAGE gels and probed with polyclonal rabbit anti-eGFP antibody to confirm expression of the full-length transgene.

### *Mass spectrometric analysis of Piwi interaction partners*

Lysis and clarification of ovary samples were performed as described above using lysis buffer with reduced detergent (0.1% Triton X-100, 0.1% Igepal). Piwi proteins with Flag, Myc, or GFP tag were purified from *Drosophila* ovaries using corresponding antibodies covalently coupled to M-270 epoxy Dynabeads (Invitrogen) (Cristea et al. 2005). Immunoprecipitation of free GFP from GFP-expressing ovaries was used as a negative control. Immunoprecipitations were performed in the presence or absence of RNase A (100  $\mu$ g/mL; 30 min at 25°C). Piwi and copurified interacting proteins were resolved on NuPAGE Novex 4%–12% Bis-Tris gels and stained with colloidal Coomassie blue. Gel fragments that contained protein bands were excised and in-gel-

trypsinized, and the peptides were extracted following the standard protocol of the Proteome Exploration Laboratory at California Institute of Technology. Peptide analyses were performed on an LTQ-FT Ultra (Thermo Fisher Scientific) equipped with a nanoelectrospray ion source (Thermo Fisher Scientific) connected to an EASY-nLC. Fractionation of peptides was performed on a 15-cm reversed-phase analytical column (75- $\mu$ m internal diameter) in-house-packed with 3- $\mu$ m C18 beads (ReproSil-Pur C18-AQ medium; Dr. Maisch GmbH). Acquired spectra were searched against the *Drosophila melanogaster* proteome using the search engine Mascot (Matrix Science, version 2.2.06), and protein inferences were performed using Scaffold (Proteome Software, version 3). For an Excel file of Piwi interaction partners, see the Supplemental Material.

### *ChIP, ChIP-seq, and RNA-seq*

ChIP was carried out using standard protocols (Moshkovich and Lei 2010). ChIP-seq and RNA-seq library construction and sequencing were carried out using standard protocols following the general principles described by Johnson et al. (2007) and Mortazavi et al. (2008), respectively. Data analysis was carried out using a combination of publicly available software tools and custom-written python scripts. Additional details regarding high-throughput data analysis are described in the Supplemental Material. For quantitative PCR (qPCR) primers, see Supplemental Table 2. GO term analysis of genes up-regulated upon Piwi knockdown was performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID) (Huang et al. 2009a,b) and FlyBase for additional assignment of GO terms. Sequencing data is available through Gene Expression Omnibus (accession no. GSE43829).

### *Antibodies*

eGFP antibody (rabbit polyclonal serum; Covance) was affinity-purified in our laboratory. Anti-myc (Millipore), anti-Flag (Sigma), Pol II (ab5408), and Pol II pSer5 (ab5131) are commercially available.

### *Imaging of ovaries*

Ovaries were fixed in 4% PFA in PBS for 20 min, permeabilized in 1% Triton X-100 in PBS, DAPI-stained (Sigma-Aldrich), washed, and mounted in 50% glycerol/PBS. Images were captured using an AxioImager microscope; an Apotome structured illumination system was used for optical sections (Carl Zeiss).

### *FLIP*

FLIP time series were captured on an LSM510 confocal microscope equipped with a 40 $\times$ /0.9 NA Imm Corr multi-immersion objective. Ovaries were dissected into halocarbon 700 oil (Sigma) and mounted under a 0.17-mm coverslip (Carl Zeiss) immediately before imaging. Two initial baseline images were captured, followed by 80–100 iterations consisting of two bleach iterations at 100% laser power (488 nm or 543 nm for GFP- and RFP-tagged proteins, respectively), followed by two images with reduced illumination intensity. FLIP series were cropped and median-filtered with a 2-pixel radius to reduce noise using FIJI (Schindelin et al. 2012) and the “Rigid Body” function of the StackReg plugin (Thévenaz et al. 1998) to correct drift when needed. Using Matlab software (The Mathworks), images were background-subtracted and corrected for acquisition bleaching. A value representing the true loss of intensity relative to the initial prebleach images, where 0 indicates no change in

intensity and 1 represents complete photobleaching, was calculated for each pixel and each bleach/capture cycle and plotted with a color lookup table and calibration bar. Scale bars and annotations were made in Inkscape (<http://inkscape.org>).

#### Preparation of polytene squashes for immunofluorescence

Flies carrying the GFP-Piwi BAC construct were backcrossed onto the *otu*[7] and *otu*[11] background. Progeny from the cross of the two lines were grown at 18°C. Stage 7–12 egg chambers were separated and transferred to a polylysine-coated microscopic slide into PBST. From here, the “smush” protocol was followed (Johansen et al. 2009), but PFA cross-linking was reduced to 10 min. Slides were imaged using an AxioImager microscope and a 63× oil immersion objective (Carl Zeiss).

#### Acknowledgments

We are grateful to Evelyn Stuwe from the Aravin laboratory for purifying the GFP antibody; I. Antoshechkin of the Millard and Muriel Jacobs Genetics and Genomics Laboratory for sequencing; D. Trout, H. Amrhein, and S. Upchurch for computational assistance; the Bloomington Stock Center for fly stocks; and S. Hess, B. Graham, and M. Sweredoski from the Proteome Exploration Laboratory at the Beckmann Institute, California Institute of Technology, for assistance with the mass spectrometry experiments. We thank members of the Aravin laboratory for critical comments on the manuscript. We thank Barbara Wold and members of the Wold laboratory for helpful discussions on ChIP protocols and analysis. A.K.R. and E.M.P. are supported by the Institutional Training Grant NIH/NRSA 5T32 GM07616, and E.M.P. is additionally supported by the Gordon Ross Medical Foundation. G.K.M. is supported by The Beckman Foundation, the Donald Bren Endowment, and NIH grant U54 HG004576. This work was supported by grants from the National Institutes of Health (R01 GM097363, R00 HD057233, and DP2 OD007371A to A.A.A.), the Searle Scholar Award (to A.A.A.), and the Ellison Medical Foundation New Scholar in Aging Award (to K.F.T.).

#### References

Agger K, Cloos P, Christensen J, Pasini D, Rose S, Rappsilber J, Issaeva I, Canaani E, Salcini A, Helin K. 2007. UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature* **449**: 731–734.

Ameyar-Zazoua M, Rachez C, Souidi M, Robin P, Fritsch L, Young R, Morozova N, Fenouil R, Descostes N, Andrau J-C et al. 2012. Argonaute proteins couple chromatin silencing to alternative splicing. *Nat Struct Mol Biol* **19**: 998–1004.

Aravin A, Hannon G, Brennecke J. 2007. The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race. *Science* **318**: 761–764.

Aravin AA, Sachidanandam R, Bourc’his D, Schaefer C, Pezic D, Toth KF, Bestor T, Hannon GJ. 2008. A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell* **31**: 785–799.

Ashe A, Sapetschnig A, Weick E-M, Mitchell J, Bagijn M, Cording A, Doebley A-L, Goldstein L, Lehrbach N, Le Pen J et al. 2012. piRNAs can trigger a multigenerational epigenetic memory in the germline of *C. elegans*. *Cell* **150**: 88–99.

Bagijn M, Goldstein L, Sapetschnig A, Weick E-M, Bouasker S, Lehrbach N, Simard M, Miska E. 2012. Function, targets, and evolution of *Caenorhabditis elegans* piRNAs. *Science* **337**: 574–578.

Brennecke J, Aravin A, Stark A, Dus M, Kellis M, Sachidanandam R, Hannon G. 2007. Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell* **128**: 1089–1103.

Brower-Toland B, Findley S, Jiang L, Liu L, Yin H, Dus M, Zhou P, Elgin S, Lin H. 2007. *Drosophila* PIWI associates with chromatin and interacts directly with HP1a. *Genes Dev* **21**: 2300–2311.

Buckley B, Burkhart K, Gu S, Spracklin G, Kershner A, Fritz H, Kimble J, Fire A, Kennedy S. 2012. A nuclear Argonaute promotes multigenerational epigenetic inheritance and germline immortality. *Nature* **489**: 447–451.

Carmell M, Girard A, van de Kant H, Bourc’his D, Bestor T, de Rooij D, Hannon G. 2007. MIWI2 is essential for spermatogenesis and repression of transposons in the mouse male germline. *Dev Cell* **12**: 503–514.

Chung W-J, Okamura K, Martin R, Lai E. 2008. Endogenous RNA interference provides a somatic defense against *Drosophila* transposons. *Curr Biol* **18**: 795–802.

Cox D, Chao A, Baker J, Chang L, Qiao D, Lin H. 1998. A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal. *Genes Dev* **12**: 3715–3727.

Cristea IM, Williams R, Chait BT, Rout MP. 2005. Fluorescent proteins as proteomic probes. *Mol Cell Proteomics* **4**: 1933–1941.

Djuranovic S, Zinchenko M, Hur J, Nahvi A, Brunelle J, Rogers E, Green R. 2010. Allosteric regulation of Argonaute proteins by miRNAs. *Nat Struct Mol Biol* **17**: 144–150.

Friedman R, Farh K, Burge C, Bartel D. 2009. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res* **19**: 92–105.

Galiana-Arnoux D, Dostert C, Schneemann A, Hoffmann J, Imler J-L. 2006. Essential function in vivo for Dicer-2 in host defense against RNA viruses in *Drosophila*. *Nat Immunol* **7**: 590–597.

Ghildiyal M, Seitz H, Horwich M, Li C, Du T, Lee S, Xu J, Kittler E, Zapp M, Weng Z et al. 2008. Endogenous siRNAs derived from transposons and mRNAs in *Drosophila* somatic cells. *Science* **320**: 1077–1081.

Grewal S, Jia S. 2007. Heterochromatin revisited. *Nat Rev Genet* **8**: 35–46.

Gunawardane L, Saito K, Nishida K, Miyoshi K, Kawamura Y, Nagami T, Siomi H, Siomi M. 2007. A slicer-mediated mechanism for repeat-associated siRNA 5′ end formation in *Drosophila*. *Science* **315**: 1587–1590.

Haase A, Fenoglio S, Muerdter F, Guzzardo P, Czech B, Pappin D, Chen C, Gordon A, Hannon G. 2010. Probing the initiation and effector phases of the somatic piRNA pathway in *Drosophila*. *Genes Dev* **24**: 2499–2504.

Handler D, Olivieri D, Novatchkova M, Gruber F, Meixner K, Mechtler K, Stark A, Sachidanandam R, Brennecke J. 2011. A systematic analysis of *Drosophila* TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors. *EMBO J* **30**: 3977–3993.

Huang DW, Sherman B, Lempicki R. 2009a. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res* **37**: 1–13.

Huang DW, Sherman B, Lempicki R. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.

Hutvagner G, Simard M. 2008. Argonaute proteins: Key players in RNA silencing. *Nat Rev Mol Cell Biol* **9**: 22–32.

Ishizu H, Nagao A, Siomi H. 2011. Gatekeepers for Piwi-piRNA complexes to enter the nucleus. *Curr Opin Genetic Dev* **21**: 484–490.

- Johansen K, Cai W, Deng H, Bao X, Zhang W, Girtton J, Johansen J. 2009. Polytene chromosome squash methods for studying transcription and epigenetic chromatin modification in *Drosophila* using antibodies. *Methods* **48**: 387–397.
- Johnson D, Mortazavi A, Myers R, Wold B. 2007. Genome-wide mapping of in vivo protein-DNA interactions. *Science* **316**: 1497–1502.
- Josse T, Teyssset L, Todeschini A-L, Sidor C, Anxolabéhère D, Ronsseray S. 2007. Telomeric *trans*-silencing: An epigenetic repression combining RNA silencing and heterochromatin formation. *PLoS Genetic* **3**: 1633–1643.
- Keller C, Adaixo R, Stunnenberg R, Woolcock KJ, Hiller S, Buhler M. 2012. HP1(Swi6) mediates the recognition and destruction of heterochromatic RNA transcripts. *Mol Cell* **47**: 215–227.
- Kiriakidou M, Tan G, Lamprinak S, De Planell-Saguer M, Nelson P, Mourelatos Z. 2007. An mRNA m7G cap binding-like motif within human Ago2 represses translation. *Cell* **129**: 1141–1151.
- Klenov M, Sokolova O, Yakushev E, Stolyarenko A, Mikhaleva E, Lavrov S, Gvozdev V. 2011. Separation of stem cell maintenance and transposon silencing functions of Piwi protein. *Proc Natl Acad Sci* **108**: 18760–18765.
- Kuramochi-Miyagawa S, Watanabe T, Gotoh K, Totoki Y, Toyoda A, Ikawa M, Asada N, Kojima K, Yamaguchi Y, Ijiri T et al. 2008. DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev* **22**: 908–917.
- Lin H, Spradling A. 1997. A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the *Drosophila* ovary. *Development* **124**: 2463–2476.
- Lin H, Yin H. 2008. A novel epigenetic mechanism in *Drosophila* somatic cells mediated by Piwi and piRNAs. *Cold Spring Harb Symp Quant Biol* **73**: 273–281.
- Maison C, Almouzni G. 2004. HP1 and the dynamics of heterochromatin maintenance. *Nat Rev Mol Cell Biol* **5**: 296–304.
- Mal'ceva N, Belyaeva E, King R, Zhimulev I. 1997. Nurse cell polytene chromosomes of *Drosophila melanogaster* *otu* mutants: Morphological changes accompanying interallelic complementation and position effect variegation. *Dev Genetic* **20**: 163–174.
- Matzke M, Aufsatz W, Kanno T, Daxinger L, Papp I, Mette M, Matzke A. 2004. Genetic analysis of RNA-mediated transcriptional gene silencing. *Biochim Biophys Acta* **1677**: 129–141.
- Mette M, Aufsatz W, van der Winden J, Matzke M, Matzke A. 2000. Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* **19**: 5194–5201.
- Mortazavi A, Williams B, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-seq. *Nat Methods* **5**: 621–628.
- Moshkovich N, Lei E. 2010. HP1 recruitment in the absence of argonaute proteins in *Drosophila*. *PLoS Genetics* **6**: e1000880.
- Muerdter F, Olovnikov I, Molaro A, Rozhkov N, Czech B, Gordon A, Hannon G, Aravin A. 2012. Production of artificial piRNAs in flies and mice. *RNA* **18**: 42–52.
- Nakayama J, Rice J, Strahl B, Allis C, Grewal S. 2001. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**: 110–113.
- Olivieri D, Sykora M, Sachidanandam R, Mechtler K, Brennecke J. 2010. An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. *EMBO J* **29**: 3301–3317.
- Onodera Y, Haag J, Ream T, Costa Nunes P, Pontes O, Pikaard C. 2005. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. *Cell* **120**: 613–622.
- Petrov DA, Aminetzach YT, Davis JC, Bensasson D, Hirsh AE. 2003. Size matters: Non-LTR retrotransposable elements and ectopic recombination in *Drosophila*. *Mol Biol Evol* **20**: 880–892.
- Pöyhönen M, de Vanssay A, Delmarre V, Hermant C, Todeschini A, Teyssset L, Ronsseray S. 2012. Homology-dependent silencing by an exogenous sequence in the *Drosophila* germline. *G3 (Bethesda)* **2**: 331–338.
- Robine N, Lau N, Balla S, Jin Z, Okamura K, Kuramochi-Miyagawa S, Blower M, Lai E. 2009. A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr Biol* **19**: 2066–2076.
- Saito K, Nishida K, Mori T, Kawamura Y, Miyoshi K, Nagami T, Siomi H, Siomi M. 2006. Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev* **20**: 2214–2222.
- Saito K, Inagaki S, Mituyama T, Kawamura Y, Ono Y, Sakota E, Kotani H, Asai K, Siomi H, Siomi M. 2009. A regulatory circuit for piwi by the large Maf gene traffic jam in *Drosophila*. *Nature* **461**: 1296–1299.
- Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B et al. 2012. Fiji: An open-source platform for biological-image analysis. *Nat Methods* **9**: 676–682.
- Shirayama M, Seth M, Lee H-C, Gu W, Ishidate T, Conte D, Mello C. 2012. piRNAs initiate an epigenetic memory of nonself RNA in the *C. elegans* germline. *Cell* **150**: 65–77.
- Sienski G, Dönertas D, Brennecke J. 2012. Transcriptional silencing of transposons by piwi and maelstrom and its impact on chromatin state and gene expression. *Cell* **151**: 964–980.
- Siomi M, Sato K, Pezic D, Aravin A. 2011. PIWI-interacting small RNAs: The vanguard of genome defence. *Nat Rev Mol Cell Biol* **12**: 246–258.
- Soppe WJ, Jasencakova Z, Houben A, Kakutani T, Meister A, Huang M, Jacobsen S, Schubert I, Fransz P. 2002. DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in *Arabidopsis*. *EMBO J* **21**: 6549–6559.
- Sugiyama T, Cam H, Verdel A, Moazed D, Grewal S. 2005. RNA-dependent RNA polymerase is an essential component of a self-enforcing loop coupling heterochromatin assembly to siRNA production. *Proc Natl Acad Sci* **102**: 152–157.
- Thévenaz P, Ruttimann U, Unser M. 1998. A pyramid approach to subpixel registration based on intensity. *IEEE Trans Image Process* **7**: 27–41.
- Vagin V, Sigova A, Li C, Seitz H, Gvozdev V, Zamore P. 2006. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science* **313**: 320–324.
- Wang S, Elgin S. 2011. *Drosophila* Piwi functions downstream of piRNA production mediating a chromatin-based transposon silencing mechanism in female germ line. *Proc Natl Acad Sci* **108**: 21164–21169.
- Wang X-H, Aliyari R, Li W-X, Li H-W, Kim K, Carthew R, Atkinson P, Ding S-W. 2006. RNA interference directs innate immunity against viruses in adult *Drosophila*. *Science* **312**: 452–454.
- Zamboni RA, Vakharia VN, Wu LP. 2006. RNAi is an antiviral immune response against a dsRNA virus in *Drosophila melanogaster*. *Cell Microbiol* **8**: 880–889.

## CHAPTER 3

### **Aub and Ago3 are recruited to nuage through two mechanisms to form the ping-pong complex assembled by Krimper**

Alexandre Webster<sup>1</sup>, Malte Wachsmuth<sup>2</sup>, Edward M. Perkins<sup>1</sup>, Ravi Sachidanandam<sup>3</sup>, & Alexei A. Aravin<sup>1, &</sup>

<sup>1</sup> Division of Biology and Bio-Engineering,  
California Institute of Technology,  
Pasadena, California 91125, USA

<sup>2</sup> Cell Biology & Biophysics Unit,  
European Molecular Biology Laboratory,  
69117 Heidelberg, Germany

<sup>3</sup> Oncological Sciences,  
Mount Sinai School of Medicine,  
New York, NY 10029, USA

& Corresponding Author: [aaa@caltech.edu](mailto:aaa@caltech.edu)

This chapter is a manuscript currently in submission for publication.

## **ABSTRACT**

**In *Drosophila*, two Piwi proteins, Aubergine (Aub) and Argonaute-3 (Ago3), localize to perinuclear ‘nuage’ granules and use guide piRNAs to target and destroy transposable element transcripts. Using static and dynamic imaging of fluorophore-tagged proteins in the *Drosophila* ovary, we show that Aub and Ago3 have distinct localization in nuage. We find that Aub and Ago3 are recruited to nuage by two different mechanisms. Aub requires a piRNA guide for recruitment to nuage, indicating that its localization depends on recognition of RNA targets. Ago3 recruitment to nuage is independent of a piRNA cargo and relies on interaction with Krimper, a stable component of nuage that is able to aggregate in the absence of other nuage proteins. We show that Krimp interacts directly with both Aub and Ago3 to coordinate the assembly of the ping-pong piRNA processing (4P) complex. Our study reveals a multi-step process responsible for the assembly and function of nuage complexes in piRNA-guided transposon repression.**

## **INTRODUCTION**

Small RNA pathways exist in many organisms and have multiple functions in sequence-specific regulation of gene expression. The three classes of small RNA are microRNA, short interfering RNA (siRNA), and Piwi-interacting RNA (piRNA). piRNA associate with Piwi clade Argonaute proteins and together repress the proliferation of transposable elements (TE) in the germline of sexually reproducing animals [81, 84]. In a Piwi-piRNA complex the small RNA is responsible for identification of target transcript and the Piwi protein fulfills the effector function [91, 92]. *D. melanogaster* has three Piwi proteins, Piwi, Aub, and Ago3, that share similar domain architectures with PAZ and Mid domains responsible for binding of guide piRNA, and a conserved DDH triad in the RNase H-like fold of the Piwi domain required for endonuclease activity [93-96]. Despite their similarity, the three Piwi proteins work non-redundantly to repress transposons [40, 97-99]. Piwi localizes to the nucleus and is involved in transcriptional repression through installation of the repressive H3K9me3 chromatin mark on target loci [100-103]. In contrast to Piwi, both Aub and Ago3 are cytoplasmic and use their



endonuclease activity to cleave transcripts complementary to their piRNA guide [38, 39, 99].

In the cell Aub and Ago3 localize to a distinct subcellular compartment termed ‘nuage’. Nuage surrounds the nucleus of nurse cells of the *Drosophila* ovary; however, it is not separated from the cytoplasm by a membrane [104, 105]. Germline granules similar to *Drosophila* nuage were first described more than 100 years ago [55, 56] and were the subject of multiple light and electron microscopy studies that showed that they are present in germ cells of almost all Metazoa [104]. However, the molecular function of nuage remained rather obscure. The recent finding that in addition to Aub and Ago3, many proteins involved in piRNA silencing, such as Vasa, Tejas, Kumo/Qin, Spindle-E, Krimper, Maelstrom, and Tudor, are localized in nuage, suggesting that nuage is a cellular compartment responsible for piRNA biogenesis and TE repression [20, 41-43, 66, 75, 106-109]. Nuage might be a location where mRNAs exiting the nucleus are scanned through a quality control process before they can pass to the cytoplasm and access ribosomes for translation. Despite the progress in understanding the composition of nuage, the role that nuage assembly plays in piRNA repression remained unknown.

Although Aub and Ago3 have very similar domain organization and subcellular localization, the two proteins play non-redundant roles in the repression of the same sets of transposons in germ cells [38, 40, 99]. Analysis of piRNAs associated with each protein revealed the basis for the unique functions of Aub and Ago3. While Aub-bound piRNA are predominantly antisense to TE sequences, Ago3-bound piRNAs are predominantly in sense orientation and therefore cannot be involved in repression directly [38, 40, 42, 99]. These and other results lead to the proposal of the so-called ping-pong model, which suggests an intimate relationship between Aub and Ago3 with cleavage products generated by one protein passed to the other [38, 99]. Specifically, cleavage of transposon mRNAs by Aub creates the 5’ end of new secondary piRNAs that are loaded into Ago3 [38, 40, 99]. Similarly, target cleavage by Ago3 can lead to loading of new antisense piRNAs into Aub, leading to amplification of piRNAs that target active TEs [39]. The Ping-pong model has very strong support from analysis of piRNA sequences in Aub and Ago3 complexes; however, the molecular mechanism describing individual

steps of ping-pong processing and the protein complexes involved in these steps remain unknown.

Using static and dynamic imaging, we found that Aub and Ago3 have different localization within nuage. Furthermore, we revealed that Aub and Ago3 are recruited to nuage through two different mechanisms. Ago3 is recruited to nuage through direct interaction with Krimper, a stable component of nuage granules that is able to form granules in the absence of other nuage components. Furthermore, Krimp coordinates assembly of the ping-pong piRNA processing (4P) complex, by directly interacting with both Ago3 and Aub. Together, our results lead to a model that explains nuage assembly in the context of the piRNA pathway and reveals molecular details of Aub and Ago3 interaction during ping-pong.

## RESULTS

### *Aub and Ago3 have distinct perinuclear localizations in nuage*

Aub and Ago3 are central components of the cytoplasmic piRNA pathway. Both proteins are non-redundantly required for repression of TEs [20, 40, 110]. Analysis of the piRNA sequences associated with Aub and Ago3 lead to the ping-pong model, which proposes that the RNA produced by Aub-dependent cleavage is transferred into unloaded Ago3 protein. However, the molecular mechanism of this transfer remains unknown. We wanted to understand the role that Aub and Ago3 play in the piRNA pathway and the mechanism of their interaction. We started by examining the subcellular localization of Aub and Ago3 in nurse cells of *Drosophila* ovaries.

Previous studies that examined the subcellular localization of Aub and Ago3 proteins using immunofluorescence microscopy led to the conclusion that both proteins are co-localized to the same subcellular compartment, perinuclear nuage granules [20, 38, 40, 99, 111]. To precisely characterize the subcellular localization of Aub and Ago3 we decided to image these proteins fused to fluorescent protein tags. We generated transgenic flies expressing Aub and Ago3 fused to GFP, and the red fluorescent protein, mKate2, and drove their expression in germ cells of fly ovary using the Maternal  $\alpha$ -Tubulin 67C (MaG4) driver [112] (Figure 1A). Tagged Aub and Ago3 are associated with piRNAs

(Figure 2A) and the profile of piRNA associated with tagged proteins was identical to those associated with endogenous proteins (data not shown). We also expressed GFP-tagged Aub under the control of its native regulatory region: this transgene rescues the sterility of *aub*<sup>HN/QC</sup> mutant females, indicating that it is fully functional. Importantly, we found that the subcellular localizations of GFP-Aub driven by the MaG4 driver or its native regulatory regions were identical in nuage and in the pole plasm, indicating that Aub localization is not affected by its overexpression when expressed by the driver (Supplemental Figure 1A). Therefore, GFP-tagged Aub and Ago3 provide us with a tool to investigate their subcellular localization with exceptional quality that is not possible with immunofluorescence methods.

Consistent with previous reports, fluorescently tagged Aub and Ago3 localize to the nuage compartment in the perinuclear region of the nurse cell cytoplasm. However, we observed important differences in the localization of Aub and Ago3 that were not noticed in previous studies. First, unlike Ago3, Aub also exhibits a diffused cytoplasmic localization. Indeed, quantification of the signals originating from the perinuclear nuage and the cytoplasm revealed that Ago3 is more enriched in nuage compared to Aub (Figure 1C). Second, we found that localization of the two proteins within nuage is different. Aub localizes around the nucleus as a smooth continuous layer with some regions of higher density. The localization of Ago3 is more punctate with the protein concentrated in several distinct granules surrounding the nucleus. Simultaneous imaging of both proteins tagged with different fluorophores showed that Ago3 granules have high concentrations of Aub protein. Conversely, nuage regions with a significant concentration of Aub might have very low Ago3 protein levels (Figure 1B). To check if fluorophore tags affect our observations, we reversed the orientation of fluorescent tags on Aub and Ago3 and observed the same characteristic localizations of Aub and Ago3 (Supplemental Figure S1B). These results indicate that Aub and Ago3 have distinct localization patterns in nuage, with Aub being distributed more uniformly and Ago3 being concentrated to granules that also contain Aub.

***Aub, but not Ago3 requires a piRNA guide for localization to nuage***

The distinct localization patterns of Aub and Ago3 in nuage suggested that different mechanisms recruit each protein to this subcellular compartment. Both Aub and Ago3 are loaded with piRNAs that recognize complementary RNA molecules. We decided to test if recognition of complementary RNA targets, which was proposed to occur in nuage, is required for recruitment of Aub and Ago3 to this compartment [20, 63]. Argonaute proteins lacking small RNA guides are not able to recognize complementary targets. To test the role of target recognition in the localization of Aub and Ago3, we generated transgenic flies that express mutant proteins (Aub-YK and Ago3-YK), which are unable to bind piRNAs due to point mutations in the conserved Y-K-K motif of the 5' piRNA binding pocket [94, 113, 114]. Purification and labeling of RNA associated with Aub-YK and Ago3-YK showed that they indeed do not associate with piRNAs (Figure 2A,B). We found that piRNA-deficient Aub was completely delocalized from nuage and was instead dispersed throughout the cytoplasm (Figure 2C). Surprisingly, localization of piRNA-deficient Ago3-YK in nuage was unperturbed. These results indicate that the mechanism of Ago3 recruitment to nuage is different from that of Aub and does not require piRNA or interaction with target RNA molecules.

To further understand the role of piRNA guides in Ago3 recruitment to nuage we used Fluorescence Recovery After Photobleaching (FRAP) to study the dynamics of wild-type and piRNA-deficient Ago3 in nuage. Perinuclear GFP-Ago3 granules were photo-bleached and the recovery of fluorescence signal was measured by time-lapse imaging. Analysis of fluorescence recovery allowed us to determine the mobile fraction of Ago3 in nuage, or percentage of protein that is free to exchange. GFP-Ago3 is dynamic in nuage with the mobile fraction being  $40.9 \pm 1.8\%$ . The absence of a piRNA guide significantly decreases the mobile fraction of Ago3-YK to  $25.2 \pm 1.5\%$  (Figure 2D, Supplemental Figure S2). This result indicates that association with piRNA enhances the mobility of Ago3 protein to exchange between nuage and cytoplasm. Therefore, the mechanisms of Aub and Ago3 recruitment to nuage seem to be completely opposite: Aub requires association with piRNA, while piRNA is dispensable for Ago3 localization. In contrast to Aub, which loses its nuage localization when devoid of piRNAs, Ago3 is becoming more restrained in nuage in the absence of a piRNA guide.

### ***Ago3 is recruited to nuage by Krimper***

The recruitment of Ago3 to nuage does not require a piRNA guide, so we sought to understand an alternative mechanism that might be guiding its recruitment. We reasoned that an interaction with one or several proteins known to localize in nuage might recruit Ago3 [20, 42, 60, 66, 106, 107, 115-117]. Since we found that Aub and Ago3 have distinct localization patterns in nuage, we examined the localization of other known nuage proteins using fluorescent protein tags (Figure 3A). Vasa and Tudor, two well-characterized nuage components, had smooth nuage localization similar to that of Aub. In contrast, a few other proteins, such as Krimp and Qin/Kumo, had more granular pattern that resembled the Ago3 distribution. We reasoned that these proteins might be responsible for the recruitment of Ago3 to nuage.

To test the role of candidate proteins in Ago3 recruitment to nuage we monitored localization of Ago3 and Aub after knocking down of Vasa, Tejas, Spindle-E, Krimper, and Tudor in germ cells. We verified the efficiency of each knock-down using two independent methods (for details see Materials and Methods). Among the proteins we targeted for depletion, only depletion of Krimp affected the localization of Ago3 but not that of Aub (Figure 3B). Krimp depletion caused mislocalization of Ago3 from perinuclear nuage and a decrease in Ago3 protein level that is particularly apparent in later stages of oogenesis. Importantly, Aub localization is not affected by Krimp depletion, supporting the conclusion that Aub and Ago3 are recruited to nuage through two different mechanisms.

Krimp was shown to be involved in piRNA-mediated repression of TEs; however, its biochemical function remained unknown [20, 66, 86, 117]. To test if Krimp is directly responsible for recruitment of Ago3 into nuage we used S2 cells that do not express Krimp, Aub and Ago3, nor the majority of other nuage components. First, we expressed each protein tagged with GFP in S2 cells and monitored their subcellular localization. In contrast to germ cells, Ago3 had a uniform cytoplasmic localization in S2 cells, indicating that it is not capable of forming granules in the absence of other nuage components or piRNA (Figure 3C). Conversely, Krimp forms prominent cytoplasmic granules when expressed in S2 cells, indicating that it is able to form granules independently of other nuage proteins and a functional piRNA pathway. It should be

noted that Krimp granules observed in S2 cells are randomly distributed in the cytoplasm and not perinuclear; therefore, they are distinct from genuine nuage formed in nurse cells of the ovary. Importantly, co-expression of Krimp with Ago3 in S2 cells sequesters Ago3 into Krimp cytoplasmic granules (Figure 3C). This result, together with the Krimp depletion experiments described above, demonstrate that Krimp is both necessary and sufficient for assembly of cytoplasmic granules and is able to recruit Ago3 into these compartments.

To further understand the role of Krimp in the formation of nuage we measured the mobility of Krimp using FRAP and compare it to the mobility of other nuage components. These experiments show that Krimp protein is extremely stable in nuage granules with only  $8.1\% \pm 2.2\%$  of the protein being mobile (Figure 3D). In contrast to Krimp, the mobile fractions of other tested proteins, Aub, Ago3, Tud, Tej, and Spn-E, exceeded 40%. This result indicates that Krimper can be considered as a structural component of nuage granules. Together our experiments implicate Krimp as a stable component of nuage that is required for recruitment of Ago3 into this subcellular compartment.

### ***Tudor domains of Krimp bind the N-terminal region of Ago3***

The ability of Krimp to form granules in S2 cells suggested that it is able to multimerize. Using co-immunoprecipitation of the Krimp protein tagged with two different tags we found that it is capable of forming dimeric or multimeric complexes (Figure 4A). When co-expressed with full length Krimp, the N-terminal fragment of Krimp (aa 1-310), which lacks any characterized domains, co-purified and co-localized in granules with full-length Krimp, indicating that this region is required for Krimp–Krimp interaction and granule formation (Figure 4A, Supplemental Figure S3).

The ability of Krimp to recruit Ago3 to granules suggests that the two proteins might interact. We examined Ago3 interactions with Krimp and other components of the piRNA pathway by co-immunoprecipitations of tagged proteins from *Drosophila* ovaries. We found that Krimp co-purifies with Ago3, while several other nuage components, Spn-E, Mael, Qin, and Vasa, did not co-purify with Ago3 under the same conditions (Figure 4B, Supplemental Figure S4A). We also found that Krimp interacts with Ago3 when

tagged proteins are transiently expressed in *Drosophila* S2 cells. This result suggests direct interaction between the two proteins, since S2 cells do not express the majority of piRNA pathway components. To further dissect the interaction between Krimp and Ago3 we mapped the interaction domains on both proteins using co-immunoprecipitation in S2 cells. The C-terminal portion of Krimp contains two putative Tudor domains. Krimp fragments that contained at least one of the two Tudor domains are able to interact with Ago3 (Figure 4C). In contrast, the N-terminal fragment of Krimp involved in its multimerization is not necessary for interaction with Ago3.

Next, we identified the region of Ago3 that interacts with Krimp. Piwi proteins contain PIWI, PAZ, and Mid domains responsible for target cleavage and binding of the 3' and 5' end of piRNA, respectively [93, 94, 96, 118]. Upstream of the PAZ domain is the N-terminal region of the protein for which the structure is not known. We found that Ago3 interacts with Krimp through its N-terminal fragment since the first 296 amino acids of Ago3 are sufficient for Krimp interaction (Figure 4D). The N-terminal fragment of Ago3 is recruited to Krimp granules when co-expressed with full length Krimp in S2 cells supporting that this fragment mediates Krimp-Ago3 interaction (Supplemental Figure S4B). Overall, we show that Krimp is able to aggregate through its N-terminal region, while its Tudor domains are responsible for binding the N-terminal region of Ago3.

To determine if Ago3 that is associated with Krimp is loaded with piRNAs we labeled small RNA isolated from Krimper-Ago3 complexes purified from ovaries. While piRNA can be easily detected in total Ago3 immunopurified from ovarian lysate, similar levels of Ago3 protein from Krimper immunoprecipitation were devoid of piRNA, indicating that Ago3 associates with Krimp in its unloaded state (Figure 4E). In contrast to Aub, the Ago3 mutant deficient in piRNA binding (Ago3-YK) still localizes to nuage (Figure 2C), further supporting the idea that piRNA loading is not required for interaction between Krimp and Ago3. Together our results suggest that Krimp-Ago3 complex is free of piRNA and loading of piRNA into Ago3 might lead to dissociation of the two proteins.

***Krimp mediates formation of a complex between Aub and Ago3 and is essential for ping-pong piRNA biogenesis***

The ping-pong model proposes an interaction between Aub and Ago3 proteins, as the Aub-piRNA produces RNA substrate for loading into to Ago3 [38, 91, 99]. As Krimp recruits Ago3 to nuage, we tested if it also interacts with Aub. First we monitored the subcellular localization of Krimp and Ago3 upon the depletion of Aub in nurse cells and found that Aub depletion resulted in delocalization of both Ago3 and Krimp from perinuclear nuage (Figure 3B). Importantly, despite delocalization from the nuclear periphery, Ago3 and Krimp co-localize in distinct granules in Aub-deficient cells. This is in contrast to the complete dispersion of Ago3 from granules observed upon Krimp depletion. These results suggest that Krimp is a primary component required for Ago3 recruitment to the granules, while Aub is necessary for attachment of Krimp/Ago3 granules to the nuclear periphery. In agreement with previous reports, we found that depletion of Ago3 did not affect Aub localization to nuage (Figure 3B) [20, 40, 117]. Taken together our results indicate that Aub functions upstream of Krimp/Ago3 and is necessary for proper recruitment of both proteins to nuage, but not for formation of granules, per se.

Next, we tested if a physical interaction between Aub and Krimp might underlie the observed genetic interactions. Co-immunoprecipitations of tagged Aub and Krimp showed that they interact when expressed in *Drosophila* ovary and S2 cells (Figure 5A, B). We further dissected the domains responsible for the interaction between each protein. Similar to the interaction between Krimp and Ago3, the N-terminal fragment of Aub interacts with the Tudor domains of Krimp (Figure 5B, C).

Our data show that Krimp is able to interact with both Ago3 and Aub using its Tudor domains (Figure 4C, 5B), while its N-terminal domain is responsible for aggregation (Figure 4A). Therefore, Krimp should be able to mediate formation of a Aub/Krimp/Krimp/Ago3 complex that puts Aub and Ago3 in close physical proximity. To test the ability of Krimp to mediate Aub and Ago3 interaction we co-expressed fluorescently tagged Aub and Ago3 in S2 cells. In the absence of Krimp, both Aub and Ago3 were dispersed in the cytoplasm; however, co-expression of Krimp led to recruitment of both Aub and Ago3 into distinct cytoplasmic granules (Figure 6A).



Therefore, Krimp is sufficient to mediate the co-localization and physical proximity between Aub and Ago3. To directly test if Krimp forms a complex that contain both Aub and Ago3 simultaneously we employed sequential immunoprecipitation (Figure 6B). First, Aub complexes were immunopurified from ovarian lysate of transgenic flies that expressed tagged Aub and Krimp proteins. After elution of Aub complexes, they were subjected to another purification step to recover complexes that contain Krimp. Western blotting revealed the presence of Ago3 protein in Aub-Krimp complexes, indicating that the triple complex containing all three proteins does indeed exist.

Taken together, our data suggest that Krimp mediates interactions between Aub and Ago3 that are necessary for the formation of an Aub/Ago3 complex. This complex, which contains both Aub and Ago3, could be essential for a key step of the ping-pong piRNA amplification process: loading of Ago3 with RNA produced by Aub-induced cleavage. If Krimp indeed plays this role, its deficiency should disrupt ping-pong piRNA processing. To test the effect of Krimp deficiency on the ping-pong cycle, we analyzed published piRNA profiles of Krimp, Ago3, Aub, and Piwi mutants [20, 40]. We monitored the output of ping-pong processing using two different tests. The first test measures the fraction of secondary piRNAs (determined by the presence of adenine residue at position 10 and the absence of uridine residue at position one) among all piRNA. The second test determines the fraction of piRNAs found in ping-pong pairs (Figure 6D) (for details see Materials & Methods). As expected from previous studies, Piwi mutation lead to large decrease in piRNA levels, but it did not have a significant effect on ping-pong processing as measured by these two tests (Figure 6D). In contrast to Piwi, deficiency in Aub completely disrupted ping-pong processing. Compared to Aub, Ago3 mutation had a relatively mild effect on ping-pong: in fact, the fraction of secondary piRNA and piRNA engaged in ping-pong pairs remains mostly unperturbed in Ago3 mutants. This result closely matches a previously published conclusion that heterotypic ping-pong between Aub and Ago3 observed in wild-type flies is replaced by homotypic Aub-Aub ping-pong in Ago3 mutants [40, 42]. Similar to the three Piwi mutants, Krimp mutant showed severe defects in its piRNA profile with a large reduction in the amount of piRNA against multiple TE families. Interestingly, our analysis shows that Krimp deficiency has essentially eliminated Ping-pong processing, as measured by a

decrease in the fraction of secondary piRNA and the fraction of piRNA present in ping-pong pairs. Indeed, the effect of Krimp deficiency on ping-pong processing was similar to that observed for Aub mutant and significantly stronger than the effect of Ago3 deficiency. This result suggests that Krimp is indeed required for ping-pong amplification. Furthermore, the function of Krimp is not restricted to heterotypic Aub-Ago3 ping-pong, but is also required for homotypic Aub/Aub ping-pong. Overall, the analysis of piRNA profiles supports the central role of Krimp in ping-pong amplification. Therefore, we propose to name the complex that includes Krimp, Aub, and Ago3 proteins as the Ping-pong piRNA processing (4P) complex (Figure 7).

## DISCUSSION

### ***Krimper assembles Ago3 / Aub complexes for ping-pong processing***

The interaction between Aub and Ago3 lies at the heart of the ping-pong piRNA processing model. This model was proposed based on the analysis of piRNA sequences residing in each Piwi protein complex [38, 99]. However, the molecular mechanism of how Aub and Ago3 interact to execute ping-pong remained poorly understood. The first step of the ping-pong cycle is cleavage of TE transcripts by the Aub/piRNA complex, which generates an RNA substrate that is next loaded into Ago3 to form a secondary piRNA. To achieve this, Aub and Ago3 might form a complex to facilitate the transfer of cleaved RNA from Aub to Ago3. Alternatively, intermediate protein(s) might transfer RNA after cleavage from Aub to Ago3. No direct association between Aub and Ago3 was ever reported, suggesting that if Aub and Ago3 ever form a complex, this interaction is likely transient and mediated by other protein(s).

Genetic studies identified several genes implicated in the ping-pong process; however, the molecular function of these proteins largely remained unknown [20, 42, 60, 66, 106, 109, 115-117]. Considering the large number of proteins involved in the ping-pong process, it is important to discriminate components that promote physical interactions between Aub and Ago3 from proteins working in other steps of this pathway. What criteria does a protein need to fulfill in order to promote the assembly of the ping-pong amplification complex? First, this protein has to directly interact with both Aub and Ago3. Second, it must be able to form a complex that contains Aub and Ago3

simultaneously. In this complex, Aub and Ago3 must be in a state that is compatible with progression of the ping-pong cycle (i.e., if Aub is loaded with piRNA then Ago should be unloaded). Third, mutation of the corresponding gene should lead to disruption of ping-pong piRNA processing. On a cellular level the protein is expected to colocalize with both Aub and Ago3. Finally, it might be expected to be a stable component of nuage that is capable of recruiting one or both Piwi proteins to this compartment.

According to our results, Krimp fulfills all the criteria for a factor that promotes the assembly of the ping-pong piRNA processing (4P) complex. Indeed, Krimp directly interacts with both Aub and Ago3. Due to the ability of Krimp to aggregate, Krimp dimers can interact with Aub and Ago3 at the same time, allowing for the formation of a Aub/Krimp/Krimp/Ago3 complex. Ago3 associates with Krimp in its unloaded form, indicating that it exists in a state ready to receive a substrate from Aub cleavage. Krimp mutation disrupts heterotypic (Aub-Ago3) and homotypic (Aub-Aub) ping-pong. Krimp co-localizes with both Aub and Ago3 and has low mobility in nuage, indicating it is a structural component of nuage. Finally, Krimp is also able to form granules by itself in the absence of other nuage proteins and is necessary and sufficient to recruit Aub and Ago3 to these granules. Overall, our data indicate that Krimp is a stable component of nuage that recruits unloaded Ago3 into this cellular compartment. Krimp interacting with Aub assembles the 4P complex that coordinates the passage of cleaved RNA fragments from Aub to unloaded Ago3, defining the essential step in the ping-pong cycle (Figure 7).

Vasa was recently proposed to coordinate the assembly of the ping-pong (Amplifier) complex [43]. However, according to published data Vasa does not fulfill most criteria required for a protein that is directly responsible for the assembly of such complex, while other criteria have not been tested. Interaction between Vasa and Ago3 was only detected in the context of a point mutant of Vasa that lacks the RNA helicase function, suggesting that this interaction is normally very transient. More importantly, the association of Vasa with the second ping-pong partner, Siwi (the silkworm homolog of Aub), appears indirect and mediated by a RNA molecule. This implies that Vasa is not involved in direct assembly of the complex containing both Aub and Ago3. Furthermore, the functional significance of the Vasa-Ago3 interaction for ping-pong processing was not tested and its ability to recruit either Piwi protein to nuage was not addressed. All these results argue

against Vasa as a component that coordinates the assembly of a 4P (Amplifier) complex. It is likely that ping-pong piRNA processing consists of many consecutive steps and mutation in components might lead to failure of the complex to progress to the next step. This would result in an arrested complex that identifies important intermediate of the pathway; however, it does not indicate that the mutated component coordinates the assembly of the complex. In fact, it might be doing exactly the opposite by coordinating the disassembly of the transient complex to allow progression towards the next processing step. It is very well possible that Vasa plays such a role in ping-pong processing; for example, as an RNA helicase it might be required for unwinding piRNA and its target RNA substrate after cleavage by Aub.

Another protein that was reported to be involved in the interaction between Aub and Ago3, Qin/Kumo better fulfills criteria for a component that is responsible for assembly of the ping-pong complex [42]. Indeed, mutation of Qin prevents the formation of a complex that contains Aub and Ago3 and also disrupts ping-pong processing. In the future, it will be important to determine if Krimp and Qin cooperate to promote the assembly of the ping-pong complex or work in two parallel pathways.

### ***A model for nuage assembly and function in the piRNA pathway***

Previous studies showed that the two central components of the cytoplasmic piRNA pathway, Aub and Ago3, along with several other proteins that are required for piRNA repression, are localized to a distinct subcellular compartment, perinuclear nuage granules. Furthermore, several mutations that disrupt nuage also cause failure in piRNA biogenesis and derepression of TEs. These observations led to the hypothesis that nuage is the subcellular compartment, where piRNA processing, target recognition, and repression take place, although this was never explicitly shown. Our results show that nuage is not one homogeneous compartment, but can be further subdivided into two structurally and functionally different sub-compartments, one containing Aub alone and another where Aub and Ago3 co-localize and likely interact. Aub covers the entire nuclear periphery, the compartment we call Aub-nuage, while Ago3 localizes to a few distinct granules embedded in the smooth Aub-nuage. This sub-compartmentalization of components in RNP granules seems to be a conserved feature of the piRNA pathway

between flies and mammals [119-121]. The distinct localization of Aub and Ago3 in nuage correlates with their different functions in the piRNA pathway.

The mechanism of nuage assembly and recruitment of Piwi proteins to this structure remained unknown. Surprisingly we found that Aub and Ago3 are recruited to nuage through different molecular mechanisms. Our data suggest that the main factor determining Aub localization to nuage is its ability to bind piRNA. Indeed, an Aub mutant that is not loaded with piRNA is completely delocalized from nuage and instead is diffused in the cytoplasm. This effect helps to explain the results of previous studies, which found that deficiencies in Rhino (Rhi), UAP56, and Cutoff (Cuff) cause delocalization of Aub from nuage [63-65]. Rhi, Cuff, and UAP56 are nuclear, precluding their direct interaction with Aub in the cytoplasm; however, as they are required for early steps of piRNA biogenesis, Aub remains unloaded and therefore mislocalizes in these mutants. The strong dependence of Aub localization on piRNA binding suggests that recognition of target RNA molecules might tether Aub in the perinuclear compartment. Alternatively, binding of piRNA might lead to conformational changes in Aub that promote its interaction with other proteins in nuage. Aub is responsible for cleavage of transposon transcripts, and to efficiently achieve this goal Aub must scan all RNA transcripts exiting the nucleus for complementarity to its piRNA guide. That explains why the smooth Aub-nuage surrounds the entire nucleus. Observed shuttling of Aub between nuage and a free cytoplasmic pool could allow Aub loaded with different piRNAs to access the nuage and scan newly exported RNAs for transposon targets. Overall, we propose that smooth Aub nuage is a compartment where cleavage of TE transcripts by the Aub/piRNA complex takes place. It is important to note that this step by itself should be sufficient to repress TE without the need of the ping-pong mechanism.

The function of the ping-pong cycle was proposed to adjust the population of piRNA to better target active transposable elements [91]. The majority of Aub molecules might not be involved in ping-pong interactions with Ago3, but instead dissociate from nuage after successful cleavage of targets, TE transcripts. However, some Aub complexes proceed to assemble a complex with Ago3 to form Aub-Ago3 nuage granules that occupy a territory distinct from Aub-nuage. Contrary to Aub, binding of piRNA is dispensable for recruitment of Ago3 to nuage (Figure 2C). Instead, our results suggest that Krimp

protein orchestrates recruitment of Ago3 to nuage and its subsequent interaction with Aub to form 4P complex (Figure 7). We identified Krimp as the most stable component of nuage among several proteins we tested (Figure 3D). Furthermore, Krimp can aggregate in vitro and form granules when expressed in non-germline S2 cells, indicating that it might be both necessary and sufficient for recruitment of Ago3 (Figure 5C). Assembly of the Aub/Krimp/Ago3 4P complex provides an opportunity for the ping-pong cycle to take effect. Importantly, Ago3 in complex with Krimp is devoid of piRNA cargo (Figure 4E). In light of the ping-pong model, Krimp recruits empty Ago3 to nuage in anticipation of Ago3 receiving the RNA that results from Aub target cleavage. Thus, in contrast to smooth Aub-nuage, the function of the distinct Ago3/Aub granules is to provide a subcellular compartment for ping-pong piRNA processing by tethering all players in one place. Overall, our data can be integrated into a dynamic model that explains the existence of two distinct nuage compartments, the forces that drive recruitment of Piwi proteins to nuage and the function of Aub and Ago3 at the different steps of the piRNA pathway (Figure 7).

## **ACKNOWLEDGEMENTS**

We thank Katalin Fejes Toth and members of the Aravin and Fejes Toth labs for discussion and critical review of this manuscript. The authors would like to thank Andres Collazo and the Beckman Imaging Facility at the California Institute of Technology, in addition to Peter Rapp, for assistance in developing our FRAP protocol; as well as Georgi Marinov for assistance with piRNA analysis, and Igor Antoshechkin of the Millard and Muriel Jacobs Genetics and Genomics Laboratory. We are grateful to Julius Brennecke, Ruth Lehmann, Phillip Zamore, Zhao Zhang, Toshie Kai, Paul Lasko, and Alexey Arkov for fly stocks and other reagents. This work was supported by grants from the National Institutes of Health (R01 GM097363 and DP2 OD007371A) and by the Searle Scholar and the Packard Fellowship Awards to A.A.A.

## MATERIALS AND METHODS

### *Fly Stocks*

Short hairpin RNA (shRNA) lines used for knockdown including *sh-White* (BDSC #33623), *sh-Aub* (BDSC #33728), *sh-Ago3* (BDSC #35232), and *sh-Krimp* (BDSC#37230), in addition to maternal alpha-Tubulin 67C-Gal4 drivers on chromosome two (BDSC #7062) or chromosome three (BDSC #7063) were purchased from the Bloomington Stock Center. The LacZ-Burdock sensor line was obtained from Dr. J. Brennecke. The Qin-GFP and UASP-FM-Qin flies were obtained from Dr. P. Zamore and Z. Zhang. The HA-Tud fly was obtained from Dr. R. Lehmann and Dr. A Arkov.

Two criteria were tested to ensure the efficacy of the knockdown by small hairpin (sh) induced RNAi. First, expression of the shRNA must knock down expression of the corresponding GFP-tagged protein target (Supplemental Figure 6A). Second, it should de-repress a piRNA silencing reporter that contains piRNA target site embedded in the 3' UTR of *LacZ* (Supplemental Figure 6B) [122].

### *Generation of Transgenic Fly Lines*

Transgenic protein constructs for injection were generated using the Gateway cloning system (Life Technologies). Genes were obtained by PCR of template cDNA taken from ovaries or testes of adult *Drosophila melanogaster*, Oregon R strain. Point mutations were engineered by overlap PCR and inserted in the pENTR-D-TOPO directional cloning vector (Life Technologies). Transgenes were cloned into the pUASP-Gateway-phiC31 fly injection vector derived from pCasPeR5-phiC31 vector containing GFP, mKate2, or Strep-FLAG tags. All transgenes created for this manuscript contain N-terminal tags, with the exception of SpnE and Zuc transgenes, which were C-terminally tagged. The expression of each transgene was controlled using the yeast upstream activation sequence promoter (UASp) stably crossed with a maternal  $\alpha$ -Tubulin67c-Gal4-VP16 (MaG4) driver, which showed strong expression in nurse cells starting at stages 2-4 of oogenesis onwards to the oocyte (Supplemental Figure 1A). Transgenes were generated in flies by PhiC31-mediated transformation (BestGene) using PhiC31 landing pads on either chromosome two (BDSC #9736) or chromosome three (BDSC #9750).

The GFP-Aub BAC line was generated by cloning of the *aub* genomic locus from the BAC clone BACN04M10 into the pCasPeR4 vector using restriction sites XhoI and SpeI. Bacterial recombineering (Gene Bridges Counter Selection kit) was used to insert an in-frame GFP tag in the start site of Aub.

*Cell Culture, Co-immunoprecipitation, & Western Blots:*

Schneider S2 cells were cultured in complete Schneider medium (10% heat inactivated FBS; 100 units of penicillin [Life technologies]; 100µg streptomycin [Life technologies]). Plasmids were generated using Gateway cloning technologies (Life technologies) using the Drosophila Gateway Vector Collection (DGVC) destination vectors, pAGW for GFP and pAFW for 3xFLAG tags expressed by the Actin5C promoter. Cells were transfected using TransIT-LT1 transfection reagent (Mirus biosciences) according to protocol but with 3.0µg of total plasmid, with 1.5 µg of each plasmid DNA containing 3xFLAG or GFP-tagged genes for double transfections, or 1.0 µg of each plasmid DNA for triple transfections.

For co-immunoprecipitation of proteins expressed in S2 cells, one well of a six-well tissue culture plate of transfected S2 cells were mechanically lysed in 150µL S2 Lysis buffer (20mM Tris at pH7.4, 150 KCl, 0.1% Tween-20, 0.1% Igepal, EDTA-free Complete Protease Inhibitor Cocktail [Roche]) and incubated for 20 minutes on ice in the presence of 100µg/mL RNase A. Supernatant was cleared by centrifugation at 4,000 x g for 10 minutes at 4°C. Input sample was collected from the supernatant at concentrations of 1-3 µg/µL. Anti-FLAG M2 beads (Sigma Aldrich) were blocked in 5mg/ml BSA for 10 minutes at 4°C, followed by washing in S2 lysis buffer. Beads were added to the supernatant and rotated at 4°C for 90 minutes. Beads were washed three times in PBS + 0.05% Tween-20 and eluted by boiling in reducing SDS loading buffer.

For co-immunoprecipitation of proteins from fly ovaries, 50 pairs of ovaries were taken from yeast-fed flies and lysed in 300µL NT2 buffer (50mM Tris at pH 7.4, 150 NaCl, 1mM MgCl<sub>2</sub>, 0.05% Igepal [NP-40], and one EDTA-free Complete Protease Inhibitor Cocktail tablet [Roche] per 10 mL). Supernatant was cleared by centrifugation at 4,000 x g for 10 minutes at 4°C. Input sample was collected from the supernatant, followed by splitting the sample into two aliquots treated with or without 100µg/mL RNase A for 10



minutes at 25°C. Anti-FLAG M2 beads (Sigma Aldrich) were blocked in 5 mg/ml BSA for 10 minutes at 4°C, followed by washing in NT2 buffer. Beads were added to the supernatant and rotated at 4°C for 90 minutes. Unbound fraction was collected for analysis and beads were washed three times for 5 minutes in 300µL NT2 buffer, and eluted by boiling in reducing SDS loading buffer, followed by running one third of the IP material on an 8% SDS-PAAG. Western Blots were probed with rabbit anti-GFP (Covance), mouse anti-GFP clone B-2 (Santa Cruz Biotechnology), or anti-FLAG M2 (Sigma Aldrich) antibody at 1:3,000 concentration. Anti-Vasa antibody was generously provided by Dr. Paul Lasko and anti-Krimper rabbit polyclonal antibody was generously provided by Dr. Toshie Kai. Both antibodies were used at a concentration of 1:20,000

### *Microscopy*

Ovaries were fixed in 4% PFA in PBS for 20 minutes, permeabilized in 1% Triton-X100 in PBS, and DAPI stained (Sigma-Aldrich). Ovaries were washed in PBS and mounted in Vectashield medium (Vector Labs). S2 cells were allowed to settle on coverslips treated with Poly-L-Lysine (Sigma-Aldrich). After gentle washing, cells were fixed in 0.5% PFA in PBS for 20 minutes followed by staining with DAPI (Sigma-Aldrich), washed, and mounted in Vectashield medium (Vector Labs). Images were captured using an AxioImager microscope; an Apotome structured illumination system was used for optical sections (Carl Zeiss).

### *Fluorescence Recovery After Photobleaching (FRAP)*

For each construct, at minimum 20 independent FRAP experiments were performed using ovaries expressing a single GFP-tagged transgene under the MaG4 driver. FRAP experiments were captured on a Zeiss LSM710 confocal microscope (Carl Zeiss AIM) equipped with a 25x/0.8 NA Imm. Corr. multi-immersion objective lens and operating Zeiss Zen Black software. Image acquisition for all experiments utilized an identical 488nm AOTF laser power setting of 7% to ensure laser power was not influencing measurements. PMT gain settings were variably set to accommodate the expression level of each GFP-tagged protein. Images were acquired at 256 x 256 pixel resolution at 0.07µm pixel size and scan speed of 614.4 ms per frame with 1.0 µs pixel dwell time. A single bleach region was defined for each experiment, consisting of a region of 7x7 pixels

equal to 0.49  $\mu\text{m}$  x 0.49  $\mu\text{m}$  and was bleached by a single iteration of 100% laser power from 488, 561 and 633nm wavelengths. Five initial pre-bleach images were captured prior to bleaching and 115 subsequent post-bleach images were acquired every 614.4ms to assess fluorescence recovery in the bleach zone. FIJI Is Just ImageJ (FIJI; <http://fiji.sc/>) software was used to analyze FRAP experiments. To account for background and photo-bleaching effects during acquisition, the mean intensity values from the bleach zone ( $BL$ ), the background zone ( $BG$ ), and the reference signal zone ( $REF$ ) were used to calculate the corrected BL ( $BL_{corr}$ ) for each acquisition frame using the equation:

$$BL_{corr} = \frac{[BL - BG]}{[REF - BG]}$$

$BL_{corr}$  values were normalized to the mean of five pre-bleach values, which were used to estimate 100% fluorescence intensity. Using the curve fitter module of FIJI, the normalized post-bleach data was fit to an exponential recovery model:

$$y = a(1 - e^{-bx}) + c$$

The mobile fraction was obtained by the sum of coefficients  $a$  and  $c$ , which describes the maximal extent of recovery for each experiment. The mean mobile fraction for each line was calculated from at minimum 20 replicate FRAP experiments taken from at least four animals from the same stable line.

#### *Nuage: Cytoplasm localization analysis*

The occupancy of GFP-tagged protein in either nuage versus cytoplasm was calculated by obtaining the ratio of mean signal intensities of a perinuclear nuage area ( $N$ ), an adjacent cytoplasmic area ( $C$ ) of equal area, and a background area ( $G$ ) using the following equation (See Figure 1C for example):

$$N:C = \frac{[N-G]}{[C-G]}$$

At least 80 different imaged nurse cells were analyzed to obtain our measurements. All images were acquired from fixed specimen using the Apotome structured illumination system and a 40x oil-immersion objective.

#### *piRNA isolation from immunopurified protein*

200 ovaries were taken from yeast-fed flies and mechanically disrupted in lysis buffer (20mM HEPES at pH 7.0, 150mM KCl, 2.5mM MgCl<sub>2</sub>, 0.5% Triton X-100, 0.5% Igepal, 100 U/mL RNasin [Promega], and one EDTA-free Complete Protease Inhibitor Cocktail [Roche] per 10mL) and supernatant clarified by centrifugation at 15,000x g. Supernatant was incubated with anti-eGFP polyclonal antibody (Covance) conjugated to Protein-G Dynabeads at 4°C and washed three times in NT2 buffer (see: *Co-immunoprecipitation & Western Blots*). Beads were spiked with 5 pmol of synthesized 42-nt (42M) RNA oligomer to assess purification efficiency (see: *Semi-Quantitative piRNA Binding Analysis*), proteinase K-digested, and phenol-extracted. Isolated RNA was CIP-treated, radiolabeled using PNK and gamma-P32-labeled ATP, and run on a 15% urea-PAAG gel.

#### *Quantification of associated piRNA from immunopurified protein*

To obtain a measure of piRNA levels co-purifying with immunoprecipitated proteins, a semi-quantitative approach was used. Gel band intensities corresponding to 29-19nt piRNA size range from autoradiographs ( $I_{piRNA}$ ) were normalized to western blot band intensities corresponding to GFP-tagged transgenes after IP ( $I_{protein}$ ). To account for potential losses in piRNA yield due to RNA purification, IP were spiked with 5pmol of synthesized 42-nt (42M) RNA oligomer prior to proteinase K digestion and phenol extracton. Gel bands corresponding to radiolabeled RNA intensities from the 42M oligomer ( $I_{42M}$ ) were subtracted from western blot band intensities. Background signal intensities from adjacent lane regions of equal area were subtracted from each band intensity measurement. These values were then used to obtain the piRNA normalized signal intensity using the following calculation:

$$piRNA \text{ normalized signal intensity}$$

$$= \frac{[I_{piRNA} - I_{BGpiRNA}]}{[I_{protein} - I_{BGprotein}] - [I_{42M} - I_{BG42M}]}$$

Band intensities were obtained using FIJI software and values graphed using Microsoft Excel.

#### *Sequence analysis of total piRNA libraries*

Sequenced total small RNA libraries from Aub, Krimp and Piwi mutant and heterozygous *Drosophila* ovaries published in Malone et al. and from Ago3 mutant and heterozygous ovaries published in Li et al. were obtained from the NCBI GEO database [20, 40]. To obtain repeat sequences, libraries were filtered for reads 23-28 nt in length and mapping to annotated repeat sequences. The read counts for repeat sequences were normalized to annotated miRNA sequences 22 nt in length. Total small RNA libraries from Malone et al. include miRNA, siRNA and piRNA sequences within the size range between 19-30nt in length. However, libraries from Li et al. treated by periodate oxidation that depleted miRNA sequences could not be normalized and were excluded from repeat sequence analyses. Mean repeats for each TE family were normalized to miRNA in heterozygous libraries include only Aub, Krimp, and Piwi heterozygous libraries but not Ago3. Annotated TE families with a mean read count in Heterozygote libraries in the lowest 10<sup>th</sup> percentile, in addition to reads mapping to Long Terminal Repeat (LTR) sequences were excluded from analysis. The fold-change in read count for each TE family was calculated by obtaining the base 2 logarithm ratio of reads in mutant (Mut.) and heterozygous (Het) libraries.

Read counts corresponding to secondary piRNA were obtained by filtering repeat sequences 23-28 nt in length for reads that were of sense orientation to annotated repeats and contained adenine (A) at position 10 and not uridine (U) at position 1. Secondary piRNA read counts for each TE family is shown as a fraction of sequences 23-28 nt in length mapping to the sense orientation of annotated repeats. The fold-change in read count for each TE family was calculated by obtaining the base 2 logarithm ratio of reads in mutant (Mut.) and heterozygous (Het) libraries.

#### *Sequence analysis of total piRNA libraries*

The normalized average fraction of reads with a corresponding ping-pong partner for each type of transposons was obtained by multiple subsampling of each of the two libraries in a pair, where the sample size was equal to the the number of reads in the library with fewest reads, or  $\min(|R1|; |R2|)$ . This was done in order to control for the larger fraction of ping-pong pairs that is expected to be observed in a larger set of reads purely by chance. For two sets of reads R1 and R2, we subsampled reads a total of N times (in this case, N=1000), and calculated the fraction of reads  $f_i$  participating in ping-pong pairs for each subsampling, i. We then defined the probabilities of a given read being in a ping-pong pair as the average fraction of reads in pairs from each subsampling:

$$\hat{p} = \frac{1}{N} \sum_{i=1}^N f_i$$

We have two such probabilities,  $\hat{p}_{RE1}$  and  $\hat{p}_{RE2}$ , for each type of repeats RE in each sample. We determine whether the observed differences between two libraries is significant, and obtain confidence bounds by estimating the difference between  $p_{RE1}$  and  $p_{RE2}$ .

For testing the null hypothesis  $H_0 : p_{RE1} = p_{RE2}$ , we calculate the following score:

$$Z = \frac{\hat{p}_{RE1} - \hat{p}_{RE2}}{\sqrt{q(1-q) \left( \frac{1}{|R_{RE1}|} + \frac{1}{|R_{RE2}|} \right)}}$$

where

$$q = \frac{\hat{p}_{RE1}|R_{RE1}| + \hat{p}_{RE2}|R_{RE2}|}{|R_{RE1}| + |R_{RE2}|}$$

The score follows a normal Z distribution with  $m = 0$  and  $s = 1$  and a p-value can be obtained from it. As we are testing more than one repeat at a time, we apply a multiple hypothesis testing correction. In this case, the Bonferroni-corrected ( $\alpha/|RE|$ ) two-tailed p-value is calculated as:

$$p_i = \min \left( 1, \frac{2 * cdf(\mathcal{N}(Z; \mu = 0; \sigma = 1))}{|RE|} \right)$$

The confidence interval of the difference  $|p_{RE1} - p_{RE2}|$  is calculated as follows:

$$\hat{p}_{RE1} - \hat{p}_{RE2} \pm Z_{\alpha} \sqrt{q(1-q) \left( \frac{1}{|R_{RE1}|} + \frac{1}{|R_{RE2}|} \right)}$$

For the 95% confidence interval,  $Z_{\alpha=0.05} = 1.96$ .

The fold-change in read count for each TE family was calculated by obtaining the base 2 logarithm ratio of reads in mutant (Mut.) and heterozygous (Het) libraries.

## REFERENCES

1. Slotkin, R. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome*. Nature reviews. Genetics, 2007. **8**(4): p. 272-285.
2. Siomi, M., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nature reviews. Molecular cell biology, 2011. **12**(4): p. 246-258.
3. Carmell, M., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis*. Genes & development, 2002. **16**(21): p. 2733-2742.
4. Cox, D., et al., *A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal*. Genes & development, 1998. **12**(23): p. 3715-3727.
5. Kuramochi-Miyagawa, S., et al., *Mili, a mammalian member of piwi family gene, is essential for spermatogenesis*. Development (Cambridge, England), 2004. **131**(4): p. 839-849.
6. Kuramochi-Miyagawa, S., et al., *Two mouse piwi-related genes: miwi and mili*. Mechanisms of development, 2001. **108**(1-2): p. 121-133.
7. Lin, H. and A. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development (Cambridge, England), 1997. **124**(12): p. 2463-2476.
8. Houwing, S., et al., *A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish*. Cell, 2007. **129**(1): p. 69-82.
9. Song, J.-J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science (New York, N.Y.), 2004. **305**(5689): p. 1434-1437.
10. Wang, Y., et al., *Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex*. Nature, 2008. **456**(7224): p. 921-926.
11. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nature cell biology, 2009. **11**(5): p. 652-658.
12. Liu, K., et al., *Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(43): p. 18398-18403.

13. Nishida, K., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. The EMBO journal, 2009. **28**(24): p. 3820-3831.
14. Vagin, V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes & development, 2009. **23**(15): p. 1749-1762.
15. Huang, H.-Y., et al., *Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish*. The EMBO journal, 2011. **30**(16): p. 3298-3308.
16. Mathioudakis, N., et al., *The multiple Tudor domain-containing protein TDRD1 is a molecular scaffold for mouse Piwi proteins and piRNA biogenesis factors*. RNA (New York, N.Y.), 2012. **18**(11): p. 2056-2072.
17. Aravin, A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Molecular cell, 2008. **31**(6): p. 785-799.
18. Aravin, A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS genetics, 2009. **5**(12).
19. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-1103.
20. Malone, C.D., et al., *Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary*. Cell, 2009. **137**(3): p. 522-35.
21. Aravin, A., et al., *A novel class of small RNAs bind to MILI protein in mouse testes*. Nature, 2006. **442**(7099): p. 203-207.
22. Girard, A., et al., *A germline-specific class of small RNAs binds mammalian Piwi proteins*. Nature, 2006. **442**(7099): p. 199-202.
23. Grivna, S., et al., *A novel class of small RNAs in mouse spermatogenic cells*. Genes & development, 2006. **20**(13): p. 1709-1714.
24. Lau, N., et al., *Characterization of the piRNA complex from rat testes*. Science (New York, N.Y.), 2006. **313**(5785): p. 363-367.
25. Aravin, A.A., et al., *Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control*. Science, 2007. **316**(5825): p. 744-747.
26. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes & development, 2013. **27**(4): p. 390-399.



27. Sienski, G., D. Dönertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-980.
28. Ro, S., et al., *Cloning and expression profiling of testis-expressed piRNA-like RNAs*. RNA (New York, N.Y.), 2007. **13**(10): p. 1693-1702.
29. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Molecular cell, 2013. **50**(5): p. 749-761.
30. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Molecular cell, 2013. **50**(5): p. 762-777.
31. Muerdter, F., et al., *A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila*. Molecular cell, 2013. **50**(5): p. 736-748.
32. Olivieri, D., et al., *An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila*. The EMBO journal, 2010. **29**(19): p. 3301-3317.
33. Ipsaro, J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-283.
34. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-287.
35. Voigt, F., et al., *Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc*. RNA (New York, N.Y.), 2012. **18**(12): p. 2128-2134.
36. Szakmary, A., et al., *The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster*. The Journal of Cell Biology, 2009. **185**.
37. Kawaoka, S., et al., *3' end formation of PIWI-interacting RNAs in vitro*. Molecular cell, 2011.
38. Gunawardane, L.S., et al., *A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila*. Science, 2007. **315**(5818): p. 1587-90.

39. Huang, H., et al., *AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification*. J Cell Biol, 2014. **206**(2): p. 217-30.
40. Li, C., et al., *Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies*. Cell, 2009. **137**(3): p. 509-21.
41. Anand, A. and T. Kai, *The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila*. EMBO J, 2012. **31**(4): p. 870-82.
42. Zhang, Z., et al., *Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains*. Mol Cell, 2011. **44**(4): p. 572-84.
43. Xiol, J., et al., *RNA clamping by vasa assembles a piRNA amplifier complex on transposon transcripts*. Cell, 2014. **157**(7): p. 1698-711.
44. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-1299.
45. Klenov, M., et al., *Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline*. Nucleic acids research, 2007. **35**(16): p. 5430-5438.
46. Chambeyron, S., et al., *piRNA-mediated nuclear accumulation of retrotransposon transcripts in the Drosophila female germline*. Proceedings of the National Academy of Sciences, 2008. **105**(39): p. 14964-14969.
47. Shpiz, S., et al., *rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus*. Nucleic acids research, 2009. **37**(1): p. 268-278.
48. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic acids research, 2011. **39**(20): p. 8703-8711.
49. Rozhkov, N., M. Hammell, and G. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes & development, 2013. **27**(4): p. 400-412.
50. Soper, S., et al., *Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis*. Developmental cell, 2008. **15**(2): p. 285-297.
51. Song, S.U., et al., *Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the gypsy retroelement of Drosophila*. Development, 1997. **124**(14): p. 2789-98.

52. al-Mukhtar, K.A. and A.C. Webb, *An ultrastructural study of primordial germ cells, oogonia and early oocytes in Xenopus laevis*. J Embryol Exp Morphol, 1971. **26**(2): p. 195-217.
53. Eddy, E.M. and S. Ito, *Fine structural and radioautographic observations on dense perinuclear cytoplasmic material in tadpole oocytes*. J Cell Biol, 1971. **49**(1): p. 90-108.
54. Mahowald, A.P., *Polar granules of Drosophila. II. Ultrastructural changes during early embryogenesis*. J Exp Zool, 1968. **167**(2): p. 237-61.
55. Hegner, R.W., *The History of the Germ Cells in the Paedogenetic Larva of Miastor*. Science, 1912. **36**(917): p. 124-6.
56. Hegner, R.W., *The Germ Cell Determinants in the Eggs of Chrysomelid Beetles*. Science, 1911. **33**(837): p. 71-2.
57. Hathaway, D.S. and G.G. Selman, *Certain aspects of cell lineage and morphogenesis studied in embryos of Drosophila melanogaster with an ultra-violet micro-beam*. J Embryol Exp Morphol, 1961. **9**: p. 310-25.
58. Okada, M., I.A. Kleinman, and H.A. Schneiderman, *Restoration of fertility in sterilized Drosophila eggs by transplantation of polar cytoplasm*. Dev Biol, 1974. **37**(1): p. 43-54.
59. Aravin, A. and D. Chan, *piRNAs meet mitochondria*. Developmental cell, 2011. **20**(3): p. 287-288.
60. Ipsaro, J.J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-83.
61. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Mol Cell, 2013. **50**(5): p. 749-61.
62. Ponting, C.P., *Tudor domains in proteins that interact with RNA*. Trends Biochem Sci, 1997. **22**(2): p. 51-2.
63. Klattenhoff, C., et al., *The Drosophila HPI homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters*. Cell, 2009. **138**(6): p. 1137-49.
64. Pane, A., et al., *The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline*. EMBO J, 2011. **30**(22): p. 4601-15.

65. Zhang, F., et al., *UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery*. Cell, 2012. **151**(4): p. 871-84.
66. Lim, A.K. and T. Kai, *Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster*. Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6714-9.
67. Patil, V.S. and T. Kai, *Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas*. Curr Biol, 2010. **20**(8): p. 724-30.
68. Anne, J. and B.M. Mechler, *Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuleen*. Development, 2005. **132**(9): p. 2167-77.
69. Kirino, Y., et al., *Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization*. RNA, 2010. **16**(1): p. 70-8.
70. Liu, H., et al., *Structural basis for methylarginine-dependent recognition of Aubergine by Tudor*. Genes Dev, 2010. **24**(17): p. 1876-81.
71. Kirino, Y., et al., *Arginine methylation of vasa protein is conserved across phyla*. J Biol Chem, 2010. **285**(11): p. 8148-54.
72. Liu, L., et al., *PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition*. Development, 2011. **138**(9): p. 1863-73.
73. Hay, B., et al., *Identification of a component of Drosophila polar granules*. Development, 1988. **103**(4): p. 625-40.
74. Anne, J., *Targeting and anchoring Tudor in the pole plasm of the Drosophila oocyte*. PLoS One, 2010. **5**(12): p. e14362.
75. Arkov, A.L., et al., *The role of Tudor domains in germline development and polar granule architecture*. Development, 2006. **133**(20): p. 4053-62.
76. Anne, J., et al., *Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila*. Development, 2007. **134**(1): p. 137-46.
77. Thomson, T. and P. Lasko, *Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning*. Genesis, 2004. **40**(3): p. 164-70.

78. Kugler, J.M. and P. Lasko, *Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis*. Fly (Austin), 2009. **3**(1): p. 15-28.
79. Liang, L., W. Diehl-Jones, and P. Lasko, *Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities*. Development, 1994. **120**(5): p. 1201-11.
80. Megosh, H.B., et al., *The role of PIWI and the miRNA machinery in Drosophila germline determination*. Curr Biol, 2006. **16**(19): p. 1884-94.
81. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline*. Curr Biol, 2001. **11**(13): p. 1017-27.
82. Bozzetti, M.P., et al., *The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2*. Proc Natl Acad Sci U S A, 1995. **92**(13): p. 6067-71.
83. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-760.
84. Vagin, V.V., et al., *A distinct small RNA pathway silences selfish genetic elements in the germline*. Science, 2006. **313**(5785): p. 320-4.
85. Aravin, A.A., et al., *Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line*. Mol Cell Biol, 2004. **24**(15): p. 6742-50.
86. Handler, D., et al., *A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors*. EMBO J, 2011. **30**(19): p. 3977-93.
87. Stapleton, W., S. Das, and B.D. McKee, *A role of the Drosophila homeless gene in repression of Stellate in male meiosis*. Chromosoma, 2001. **110**(3): p. 228-40.
88. Kibanov, M.V., et al., *A novel organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing*. Mol Biol Cell, 2011. **22**(18): p. 3410-9.

89. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(46): p. 18760-5.
90. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-9.
91. Aravin, A.A., G.J. Hannon, and J. Brennecke, *The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race*. Science, 2007. **318**(5851): p. 761-4.
92. Siomi, M.C., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nat Rev Mol Cell Biol, 2011. **12**(4): p. 246-58.
93. Ma, J.B., K. Ye, and D.J. Patel, *Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain*. Nature, 2004. **429**(6989): p. 318-22.
94. Ma, J.B., et al., *Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein*. Nature, 2005. **434**(7033): p. 666-70.
95. Yuan, Y.R., et al., *Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage*. Mol Cell, 2005. **19**(3): p. 405-19.
96. Song, J.J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science, 2004. **305**(5689): p. 1434-7.
97. Lin, H. and A.C. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development, 1997. **124**(12): p. 2463-76.
98. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-60.
99. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-103.
100. Sienski, G., D. Donertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-80.

101. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes Dev, 2013. **27**(4): p. 390-9.
102. Rozhkov, N.V., M. Hammell, and G.J. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes Dev, 2013. **27**(4): p. 400-12.
103. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic Acids Res, 2011. **39**(20): p. 8703-11.
104. Eddy, E.M., *Germ plasm and the differentiation of the germ cell line*. Int Rev Cytol, 1975. **43**: p. 229-80.
105. Mahowald, A.P., *Polar granules of drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis*. J Exp Zool, 1971. **176**(3): p. 345-52.
106. Patil, V.S. and T. Kai, *Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas*. Curr Biol, 2010.
107. Findley, S.D., et al., *Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage*. Development, 2003. **130**(5): p. 859-71.
108. Lim, A.K., L. Tao, and T. Kai, *piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline*. J Cell Biol, 2009. **186**(3): p. 333-42.
109. Nishida, K.M., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. EMBO J, 2009. **28**(24): p. 3820-31.
110. Schupbach, T. and E. Wieschaus, *Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology*. Genetics, 1991. **129**(4): p. 1119-36.
111. Harris, A.N. and P.M. Macdonald, *Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C*. Development, 2001. **128**(14): p. 2823-32.
112. Bossing, T., C.S. Barros, and A.H. Brand, *Rapid tissue-specific expression assay in living embryos*. Genesis, 2002. **34**(1-2): p. 123-6.
113. Djuranovic, S., et al., *Allosteric regulation of Argonaute proteins by miRNAs*. Nat Struct Mol Biol, 2010. **17**(2): p. 144-50.

114. Parker, J.S., S.M. Roe, and D. Barford, *Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex*. Nature, 2005. **434**(7033): p. 663-6.
115. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-7.
116. Snee, M.J. and P.M. Macdonald, *Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components*. J Cell Sci, 2004. **117**(Pt 10): p. 2109-20.
117. Nagao, A., et al., *Gender-Specific Hierarchy in Nuage Localization of PIWI-Interacting RNA Factors in Drosophila*. Front Genet, 2011. **2**: p. 55.
118. Wang, Y., et al., *Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes*. Nature, 2009. **461**(7265): p. 754-61.
119. Aravin, A.A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS Genet, 2009. **5**(12): p. e1000764.
120. Aravin, A.A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Mol Cell, 2008. **31**(6): p. 785-99.
121. Shoji, M., et al., *The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline*. Dev Cell, 2009. **17**(6): p. 775-87.
122. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Mol Cell, 2013. **50**(5): p. 762-77.
123. Vermaak, D., S. Henikoff, and H.S. Malik, *Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila*. PLoS Genet, 2005. **1**(1): p. 96-108.
124. Obbard, D.J., et al., *The evolution of RNAi as a defence against viruses and transposable elements*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1513): p. 99-115.
125. Simkin, A., et al., *Recurrent and recent selective sweeps in the piRNA pathway*. Evolution, 2013. **67**(4): p. 1081-90.
126. Comeron, J.M., *A method for estimating the numbers of synonymous and nonsynonymous substitutions per site*. J Mol Evol, 1995. **41**(6): p. 1152-9.



127. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nat Cell Biol, 2009. **11**(5): p. 652-8.
128. Anne, J., *C-terminal moiety of Tudor contains its in vivo activity in Drosophila*. PLoS One, 2010. **5**(12): p. e14378.
129. Creed, T.M., et al., *Novel role of specific Tudor domains in Tudor-Aubergine protein complex assembly and distribution during Drosophila oogenesis*. Biochem Biophys Res Commun, 2010. **402**(2): p. 384-9.
130. Vagin, V.V., G.J. Hannon, and A.A. Aravin, *Arginine methylation as a molecular signature of the Piwi small RNA pathway*. Cell Cycle, 2009. **8**(24): p. 4003-4.
131. Arkov, A.L. and A. Ramos, *Building RNA-protein granules: insight from the germline*. Trends Cell Biol, 2010. **20**(8): p. 482-90.
132. Siomi, M.C., T. Mannen, and H. Siomi, *How does the royal family of Tudor rule the PIWI-interacting RNA pathway?* Genes Dev, 2010. **24**(7): p. 636-46.
133. Cavey, M., et al., *Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation*. Development, 2005. **132**(3): p. 459-68.
134. Vagin, V.V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes Dev, 2009. **23**(15): p. 1749-62.
135. Bedford, M.T. and S.G. Clarke, *Protein arginine methylation in mammals: who, what, and why*. Mol Cell, 2009. **33**(1): p. 1-13.
136. Krause, C.D., et al., *Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential*. Pharmacol Ther, 2007. **113**(1): p. 50-87.
137. Wang, J., et al., *Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis*. Curr Biol, 2009. **19**(8): p. 640-4.
138. Saxe, J.P., et al., *Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline*. EMBO J, 2013. **32**(13): p. 1869-85.
139. Pandey, R.R., et al., *Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16492-7.

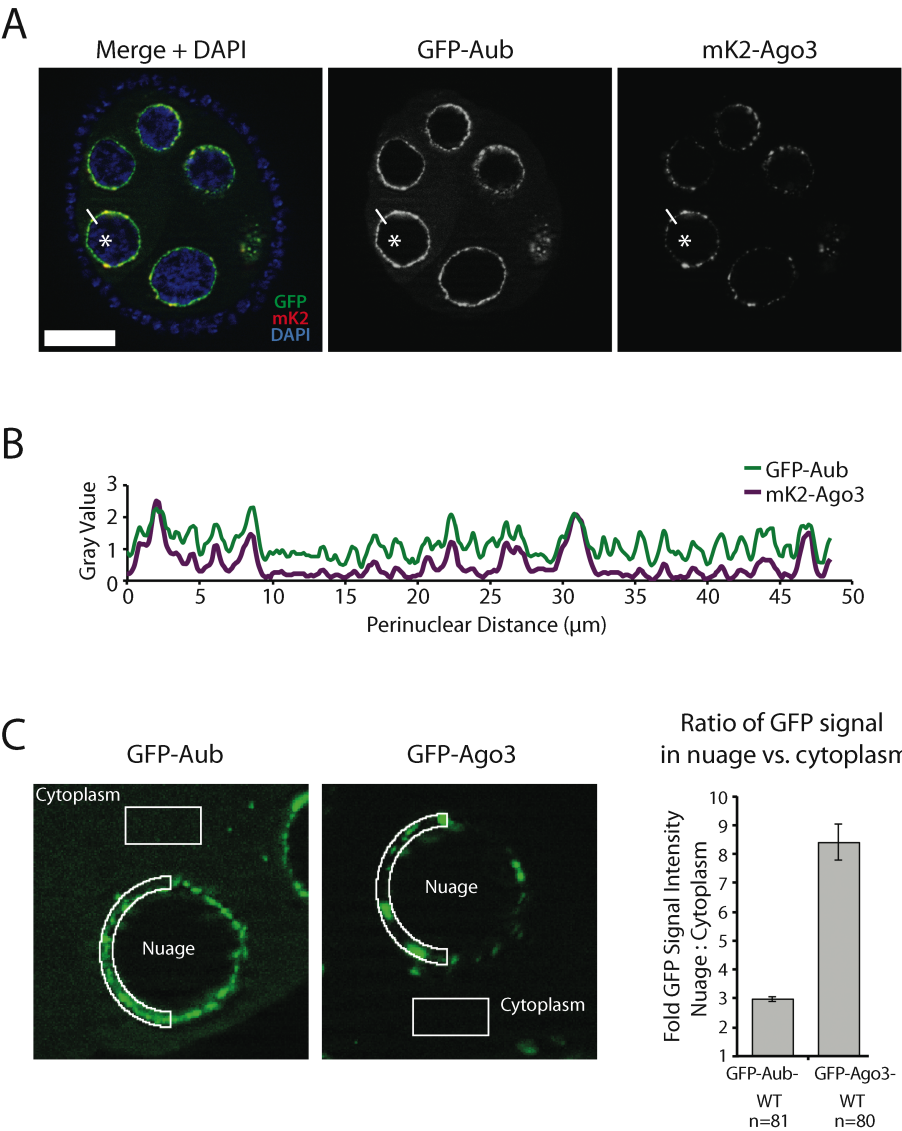
140. Cora, E., et al., *The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs*. RNA, 2014. **20**(6): p. 773-81.

## FIGURES AND FIGURE LEGENDS

### ***Figure 1. Aub and Ago3 have distinct subcellular localizations.***

(A) The localization pattern of Aub (GFP-Aub, green) appears as a smooth perinuclear layer, while Ago3 (mK2-Ago3, red) forms punctate granules around the nuclei. Proteins are expressed simultaneously in nurse cells of ovaries using the same driver. DNA is stained with DAPI (blue). Scale bar: 20 $\mu$ m. (B) The profile of Aub (green) and Ago3 (purple) localization around a single nucleus shows the distinct localizations of each protein. Signal intensities of Aub (GFP-Aub) and Ago3 (mK2-Ago3) were traced along the perimeter of the nurse cell nucleus marked with an asterisk in Figure 1A. The line intersecting the perinuclear region marks the start of measurements, continuing clockwise around the perimeter. (C) Ago3 is more concentrated in nuage, while Aub is more dispersed in the cytoplasm. The fluorescence signals of GFP-Aub and GFP-Ago3 were quantified in nuage and cytoplasm (regions with identical surface area shown on the figure— see materials & methods for details). Shown is the mean ratio of nuage to cytoplasmic signal measured for 81 and 80 individual nurse cells, respectively. Standard deviation from the mean is shown for each protein.

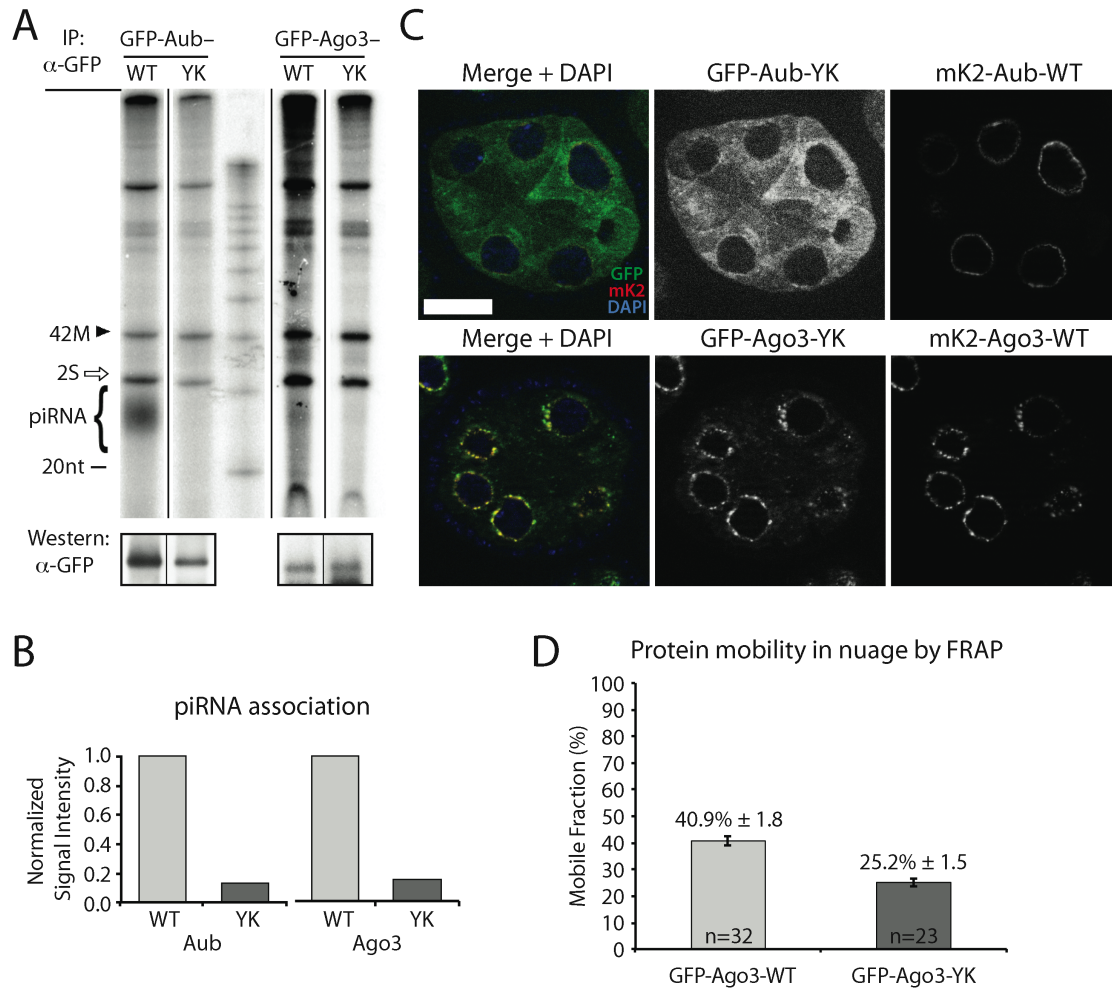
Webster et al. Figure 1



**Figure 2. *Aub* requires a piRNA guide for localization to nuage.**

(A) piRNA-deficient Aub (Aub-YK) and Ago3 (Ago3-YK) mutant proteins show a loss in associated piRNAs compared to their wild-type protein counterparts. piRNA-deficient Aub-YK or Ago3-YK mutant proteins were expressed in ovaries, immunoprecipitated, and their associated small RNA isolated and 5' radiolabeled. The strong band at 30 nt is the abundant 2S ribosomal RNA (rRNA) (white arrow). A 42 nt RNA (42M) was spiked into each immuno-purified sample to control for labeling efficiency and RNA loss during isolation. The levels of immunopurified proteins are shown by western blot (lower panels). (B) Quantification of piRNA association of wild-type and mutant Aub and Ago3 shown on panel (A). The signals of radiolabeled piRNA were normalized to the amount of immunoprecipitated protein and to the amount of spike RNA (see materials and methods for details). (C) The piRNA-binding deficient Aub-YK does not localize to nuage, but is diffused in the cytoplasm. Wild-type protein tagged with mKate2 was co-expressed in the same cells (upper panel). In contrast to Aub, the localization of piRNA-binding deficient Ago3-YK is indistinguishable from that of wild-type Ago3. DNA is stained with DAPI (blue). Scale bar: 20µm. (D) The mobility of the piRNA-binding deficient Ago3-YK mutant is decreased in nuage. The fractions of mobile wild-type and Ago3-YK proteins were measured by FRAP on 32 and 23 independent cells, respectively. Bars show standard error values.

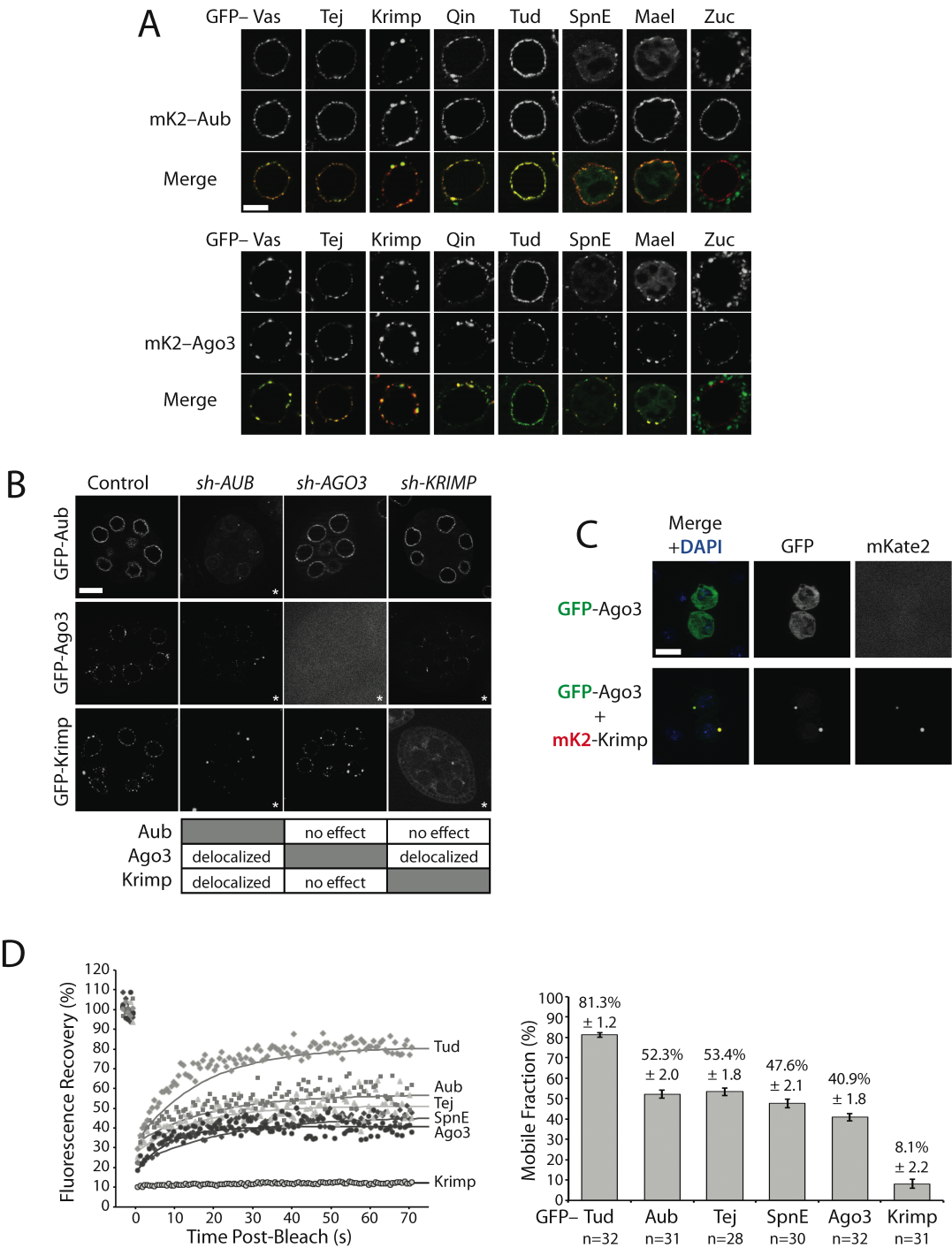
Webster et al. Figure 2



**Figure 3. *Krimper recruits Ago3 to granules.***

(A) Co-localization of Aub and Ago3 with other nuage components. Aub (mK2-Aub, red; upper panel) or Ago3 (mK2-Ago3 red; lower panel) were co-expressed in nurse cells of *Drosophila* ovaries with GFP-tagged piRNA pathway components (green) using the same driver. Vas, Tej, and Tud show strong co-localization with Aub, while Ago3 is strongly co-localized with Krimp and Qin in nuage granules. SpnE and Mael localize in nuage but also to the nucleus. Zuc shows little to no co-localization with either Aub or Ago3. Scale bar: 10µm. (B) The genetic interaction between Aub, Ago3, and Krimp. Aub, Ago3, and Krimp were knocked down in germ cells of flies expressing GFP-tagged proteins using shRNA expressed under control of the MaG4 driver. The asterisk in each image indicates perturbed localization of the corresponding protein compared to wild type. Aub is required for Ago3 and Krimp nuage localization, while Aub localizes independently of Krimp and Ago3. Krimp is required for both localization and normal protein level of Ago3, but its localization is unchanged upon Ago3 depletion. (C) Krimp recruits Ago3 into cytoplasmic granules in S2 cells. S2 cells were transfected with plasmids encoding GFP-tagged Ago3 (green) alone or in combination with mKate2-tagged Krimp protein (red). GFP-tagged Ago3 is diffused in the cytoplasm when expressed alone. The expression of Krimp induces Ago3 recruitment into Krimp cytoplasmic granules. DNA is stained with DAPI (blue). Scale bar: 10µm. (D) Krimp is the stationary component of nuage granules. FRAP shows large differences in the mobility of different proteins in nuage. All proteins were tagged with GFP and expressed in germ cells under control of the MaG4 driver. Representative FRAP experiments are shown on the left. The mobile fraction was determined by modeling the recovery to an exponential recovery curve superimposed on respective datasets. The quantification of the mobile fractions for each of the proteins is shown on the right (n indicates the number of independent experiments used for FRAP measurements). More than 80% of Tud is mobile in nuage, while the mobile fractions of Aub, Tej, SpnE, and Ago3 are between 40 and 50%. Krimp protein is least mobile, with a mobile fraction of ~8%. Bars show standard error values.

Webster et al. Figure 3

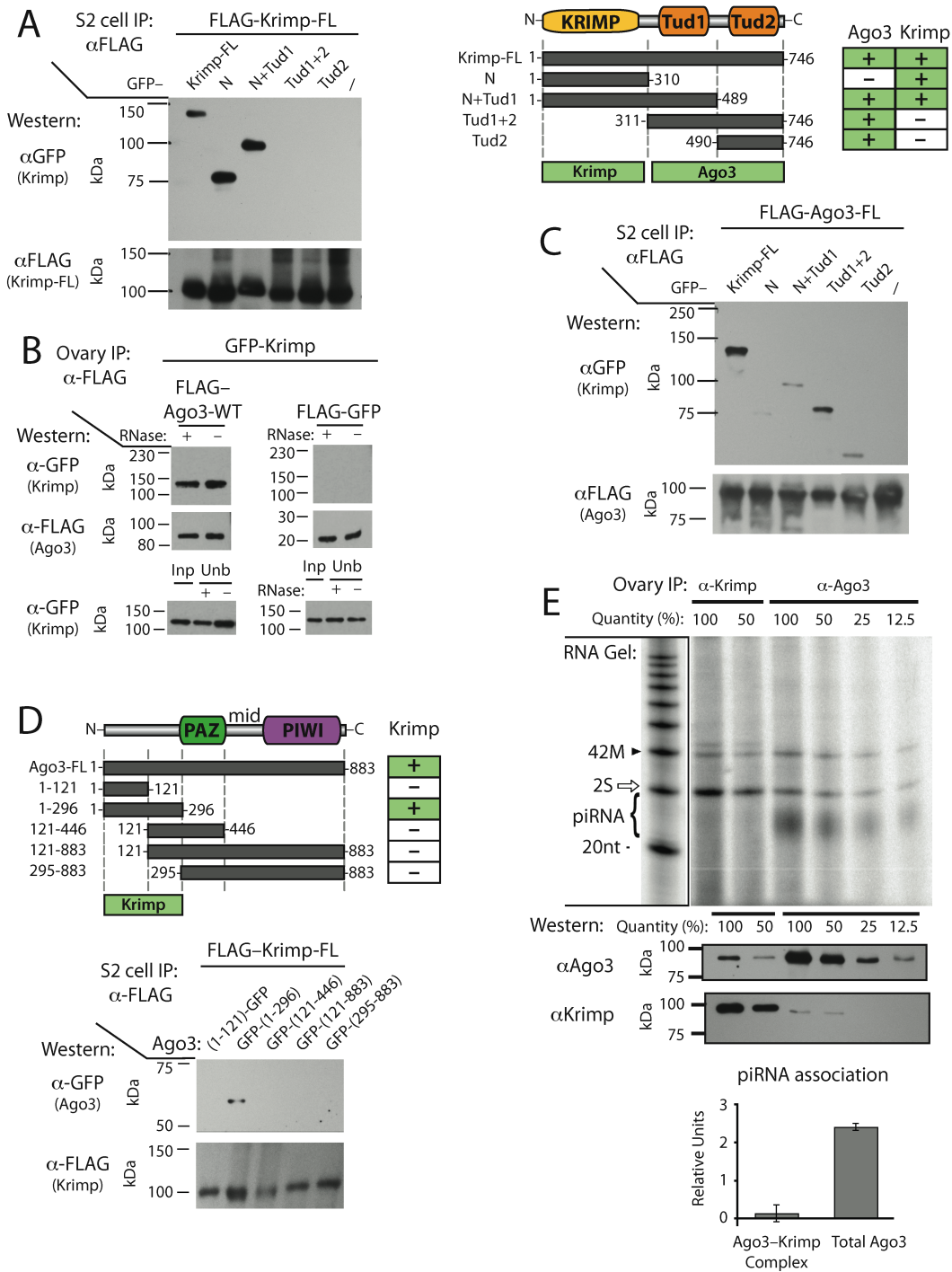




**Figure 4. Separate domains of Krimper are responsible for dimerization and association with Ago3.**

(A) The N-terminal region of Krimper is required for Krimper dimerization. FLAG-tagged full-length Krimper (Krimper-FL) was co-expressed with GFP-tagged fragments of Krimper (shown on the right panel) in S2 cells. Fragments of Krimper containing the first 310 amino acid residues co-purified with Krimper-FL, while fragments lacking N-terminal residues did not co-purify. (B) Krimper forms a protein-protein complex with Ago3 in *Drosophila* ovary. FLAG-Ago3 and GFP-Krimper were co-expressed in fly ovaries. FLAG-Ago3 or control (FLAG-tagged GFP) proteins were immunoprecipitated with anti-FLAG beads in the presence (+) or absence (-) of RNase A followed by Western blotting to detect GFP- and FLAG-tagged proteins. The interaction between Ago3 and Krimper is independent of RNA. The lower panel shows Input (Inp) and the unbound fractions (Unb). (C) Tudor domains of Krimper bind Ago3. FLAG-tagged full length Ago3 (FLAG-Ago3-FL) was co-expressed with GFP-tagged fragments of Krimper in S2 cells. FLAG-Ago3 was immunoprecipitated followed by Western to detect GFP-Krimper fragments. Ago3 co-purifies with fragments of Krimper containing either of its two Tudor domains, but not with the N-terminal fragment responsible for dimerization. (D) Krimper interacts with the N-terminal fragment of Ago3. FLAG-tagged full length Krimper was co-expressed with GFP-tagged fragments of Ago3 in S2 cells. The N-terminal Ago3 fragment containing the first 296 amino acid residues co-purifies with Krimper, while other Ago3 fragments are not able to interact with Krimper. (E) Ago3 in complex with Krimper is not bound to piRNA. Total Ago3 and Ago3-Krimper complexes were purified from *Drosophila* ovaries. Co-purified RNA was labeled and resolved on the gel (top panel), while the amount of Ago3 and Krimper proteins in immunoprecipitates was determined by Western blotting (middle panels). The amount of piRNA in each immunoprecipitate was normalized to the amount of Ago3 protein (lower panel). An unlabeled 42 nt RNA (42M) was spiked into each sample to control for labeling efficiency and RNA loss.

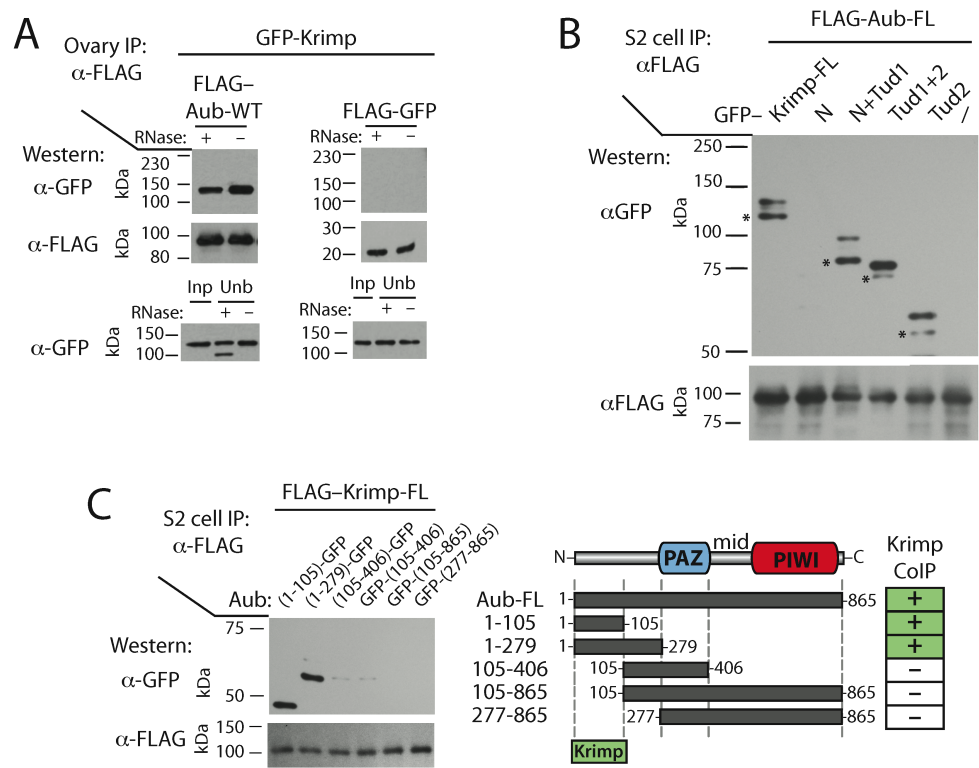
Webster et al. Figure 4



**Figure 5. *Krimper* interacts with *Aub*.**

(A) Krimp forms a complex with Aub in *Drosophila* ovary. FLAG-Aub and GFP-Krimp were co-expressed in fly ovaries. FLAG-Aub or control (FLAG-GFP) proteins were immunoprecipitated with anti-FLAG beads in the presence (+) or absence (-) of RNase A followed by Western blotting to detect GFP- and FLAG-tagged proteins. The interaction between Aub and Krimp is independent of RNA. The lower panel shows Input (Inp) and the unbound (Unb) fractions. (B) The Tudor domains of Krimp interact with Aub. FLAG-Aub was co-expressed with GFP-tagged fragments of Krimp protein in S2 cells. Krimp fragments containing either of the two Tudor domains co-purify with Aub. Asterisks indicate bands corresponding to degradation products of Krimp-GFP. (C) The N-terminal fragment of Aub binds Krimp. FLAG-Krimp was co-expressed with GFP-tagged fragments of Aub in S2 cells. The N-terminal fragment of Aub containing the first 105 amino acid residues co-purifies with Krimp, while other Aub fragments lacking this region did not co-purify.

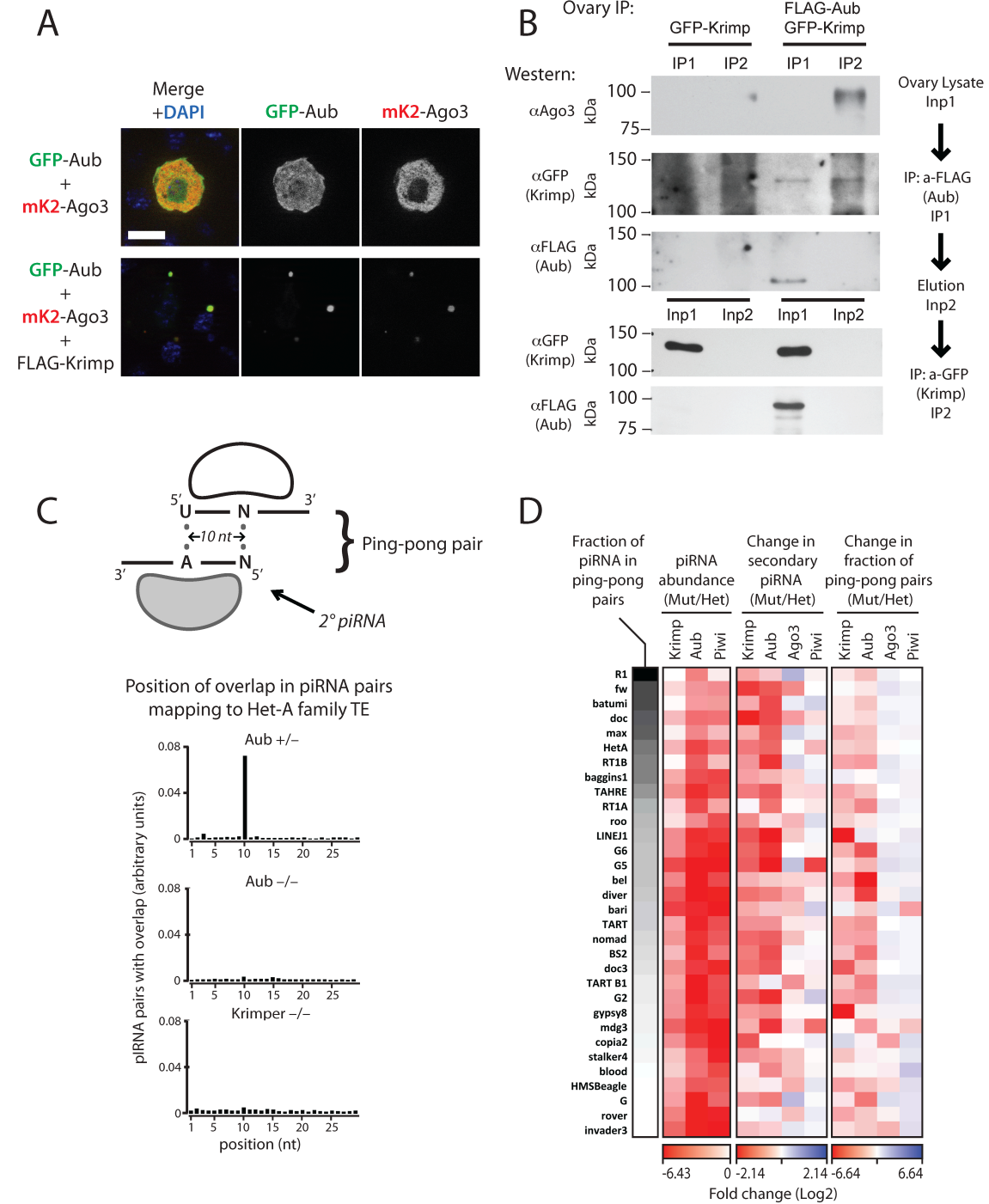
Webster et al. Figure 5



***Figure 6. Krimper assembles the complex that contains both Aub and Ago3 proteins and is required for ping-pong piRNA processing.***

(A) Krimper recruits Aub and Ago3 into cytoplasmic granules. GFP-tagged Aub and mKate2-tagged Ago3 were co-expressed in S2 cells. In the absence of Krimper, Aub and Ago3 have diffused cytoplasmic localization; however, both proteins are recruited into the same cytoplasmic granules upon Krimper expression. Scale bar: 10µm. (B) Krimper forms a triple complex with Aub and Ago3. Ovarian lysates from flies co-expressing FLAG-Aub and GFP-Krimper (and control flies expressing GFP-Krimper alone) were subjected to tandem co-immunoprecipitation. The first anti-FLAG immunoprecipitation purifies Aub, followed by elution and anti-GFP immunoprecipitation to purify Aub-Krimper complexes. Finally, the presence of Ago3 was determined by Western blotting with an antibody against the native protein. (C) The pairs of complementary piRNAs formed by ping-pong processing are separated by 10nt. Primary piRNA predominantly has a uridine residue at the first position. Corresponding secondary piRNA has an adenine residue at position 10 and lacks nucleotide bias at the first position. The graphs below show the fraction of complementary piRNA that have different overlap between their 5' ends. The predominant peak at 10nt is seen in wild-type piRNA populations indicating that a large fraction of piRNA overlap by 10 nt at their 5' ends. This peak disappears in piRNA from Aub and Krimper mutants. (D) Krimper is required for ping-pong piRNA processing. The fraction of secondary piRNAs and the fraction of the ping-pong pairs that have 10nt overlap were analyzed in Krimper, Piwi, Aub, and Ago3 mutants and control (heterozygous) flies. The changes in mutant compared to control for each family of transposable elements are shown on the heat map. Only TE families with high level of piRNA expression and an abundance in ping-pong pairs are shown. The same analysis for all TE families is shown in Supplemental Figure S5.

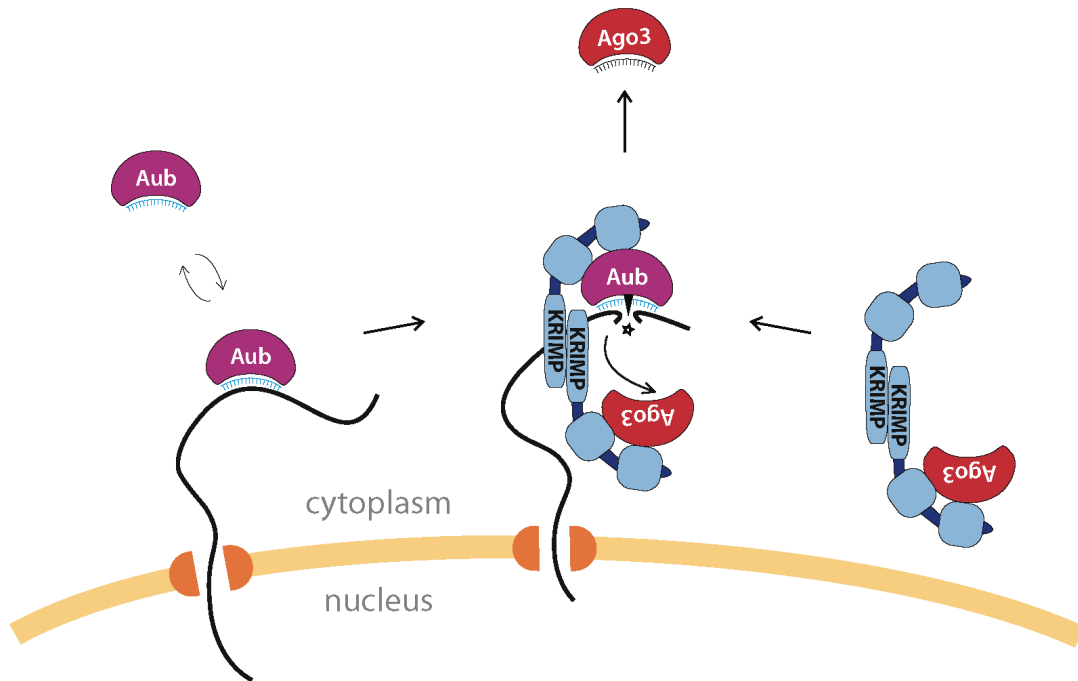
Webster et al. Figure 6



***Figure 7. A model for recruitment of Aub and Ago3 to nuage and formation of the ping-pong piRNA processing (4P) complex.***

Aub-piRNA complexes shuttle between the cytoplasm and nuage interrogating RNA transcripts as those exit the nucleus. The recognition of complementary target by piRNA might recruit Aub to nuage. Ago3 that is devoid of piRNA is recruited to nuage through the interaction with Krimp, a stable component of nuage granules. Upon recognition of complementary target by the Aub-piRNA complex, the target is either cleaved, followed by release of Aub-piRNA from nuage (1) or Aub, engaged to an RNA target, interacts with the Ago3-Krimp complex (2). In this 4P complex, the Krimp dimer places unloaded Ago3 protein in proximity to Aub associated with target RNA. Endonucleolytic cleavage of target by Aub generates the RNA substrate that is loaded into Ago3. Once bound to RNA, Ago3 may be released from the Krimper complex.

Webster et al. Figure 7

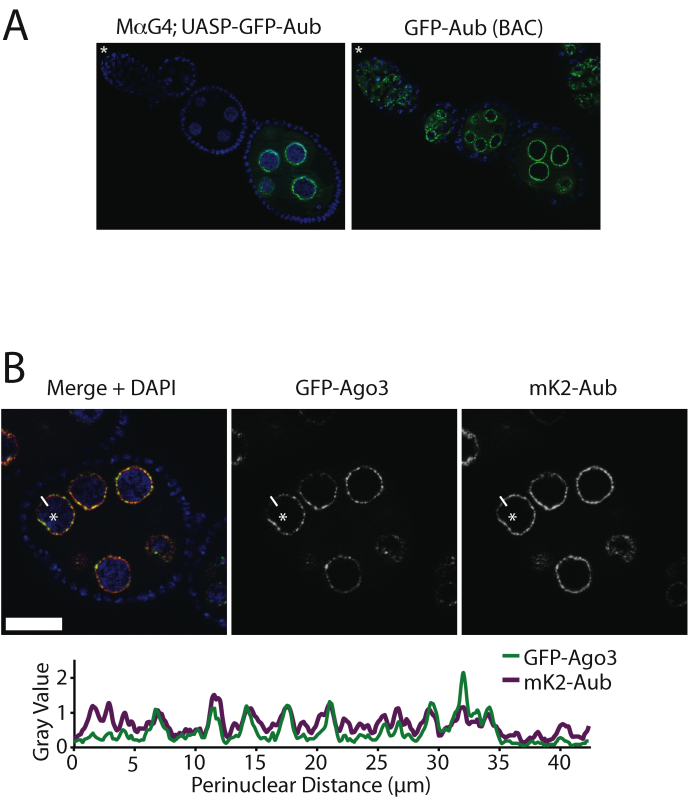




## SUPPLEMENTAL FIGURES AND FIGURE LEGENDS

### ***Supplemental Figure 1: The expression of Aub and Ago3 transgenes in Drosophila ovaries.***

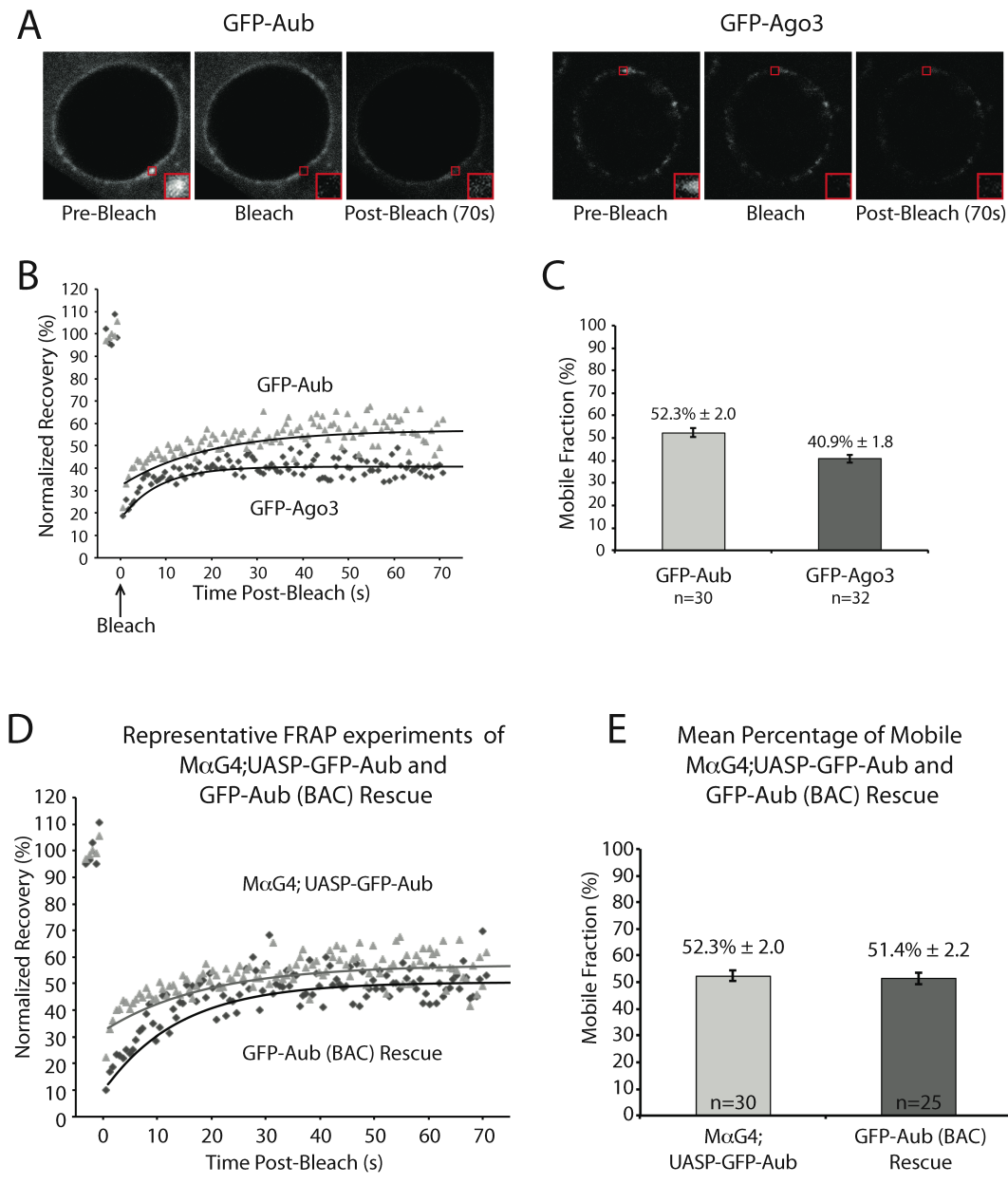
(A) Expression of MaG4 driven UASp-GFP-Aub-WT (left panel; green) begins at stage 4-6 of oogenesis, while a GFP-Aub-WT BAC line (right panel; green) expressing transgenic protein under the native Aub promoter begins at stage one of the germarium and continues onward to the oocyte. The asterisk at the top left of each image marks the anterior tip of the germarium for each line. DNA is stained with DAPI (blue). (B) The localization of tagged Aub and Ago3 proteins are independent of fluorophore tags. Perinuclear localization of Aub (purple) is different from Ago3 localization (green). In the lower panel, gray values for mKate2-Aub and GFP-Ago3 are displayed along the perimeter of the nurse cell marked with an asterisk in the upper panel. The line intersecting the perinuclear region marks the start of measurements, continuing clockwise around the perimeter. Scale bar: 20 $\mu$ m.



***Supplemental Figure 2: Dynamics of Aub and Ago3 in nuage.***

(A) Representative FRAP experiments on GFP-tagged Aub and Ago3 in nurse cells of the *Drosophila* ovary. The red box delineates the FRAP bleach region of nuage pre-bleach, immediately after bleach, and 70 seconds post-bleach. The lower right corner shows the bleach region magnified. (B) Representative FRAP experiments of Aub and Ago3 in nuage granules. The mobile fraction was determined by modeling the recovery to an exponential recovery curve (superimposed on respective data-sets). The recovery of each curve is dependent on steady-state mobility, or exchange, of GFP tagged protein within the bleached zone ( $0.49 \mu\text{m}^2$ ). The bleaching time point ( $T=0$ ) is marked by an arrow. For each experiment, the 100% Normalized Recovery limit is set to the normalized mean intensity of the five pre-bleach time points ( $T<0$ ). (C) The mean fraction of mobile Aub and Ago3 proteins of replicate FRAP experiments ( $n=30$  and  $n=32$ , respectively) shows greater mobility of Aub within nuage compared to Ago3. Bars show standard error values for each set of experiments. (D) Representative FRAP experiments of comparing UASp-GFP-Aub expressed by the MaG4 driver over a wild-type background versus a GFP-Aub BAC line expressed in an Aub-mutant background. Both representative FRAP experiments show similar levels of fluorescence recovery after photobleaching, suggesting that expression of Aub over a wild type background does not affect the dynamics of transgenic protein. The mobile fraction was determined by modeling the recovery to an exponential recovery curve (superimposed on respective data-sets). (E) FRAP measurements are not affected by expression over endogenous protein in nurse cells. The mean fraction of mobile MaG4; UASp-GFP-Aub expressed over a wild-type background versus a GFP-Aub BAC line expressed in an Aub-mutant background from replicate FRAP experiments ( $n=30$  and  $n=25$ , respectively) show similar fractions of mobile protein ( $52.3\% \pm 2.0\%$  and  $51.4\% \pm 2.2\%$ , respectively).

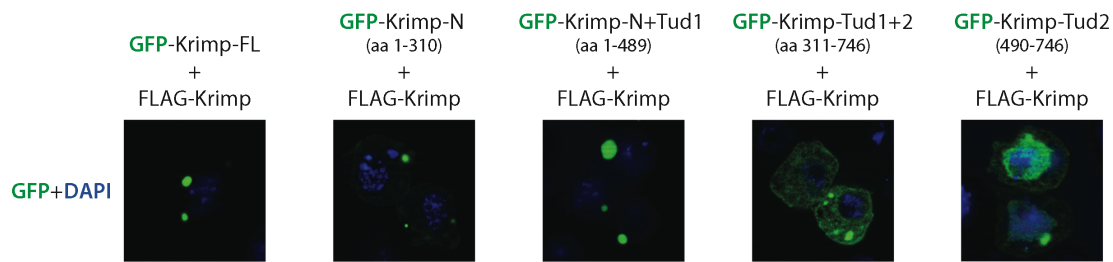
Webster et al. Supplemental Figure 2



***Supplemental Figure 3: Krimper multimerization is mediated by self-interactions in the N-terminal region.***

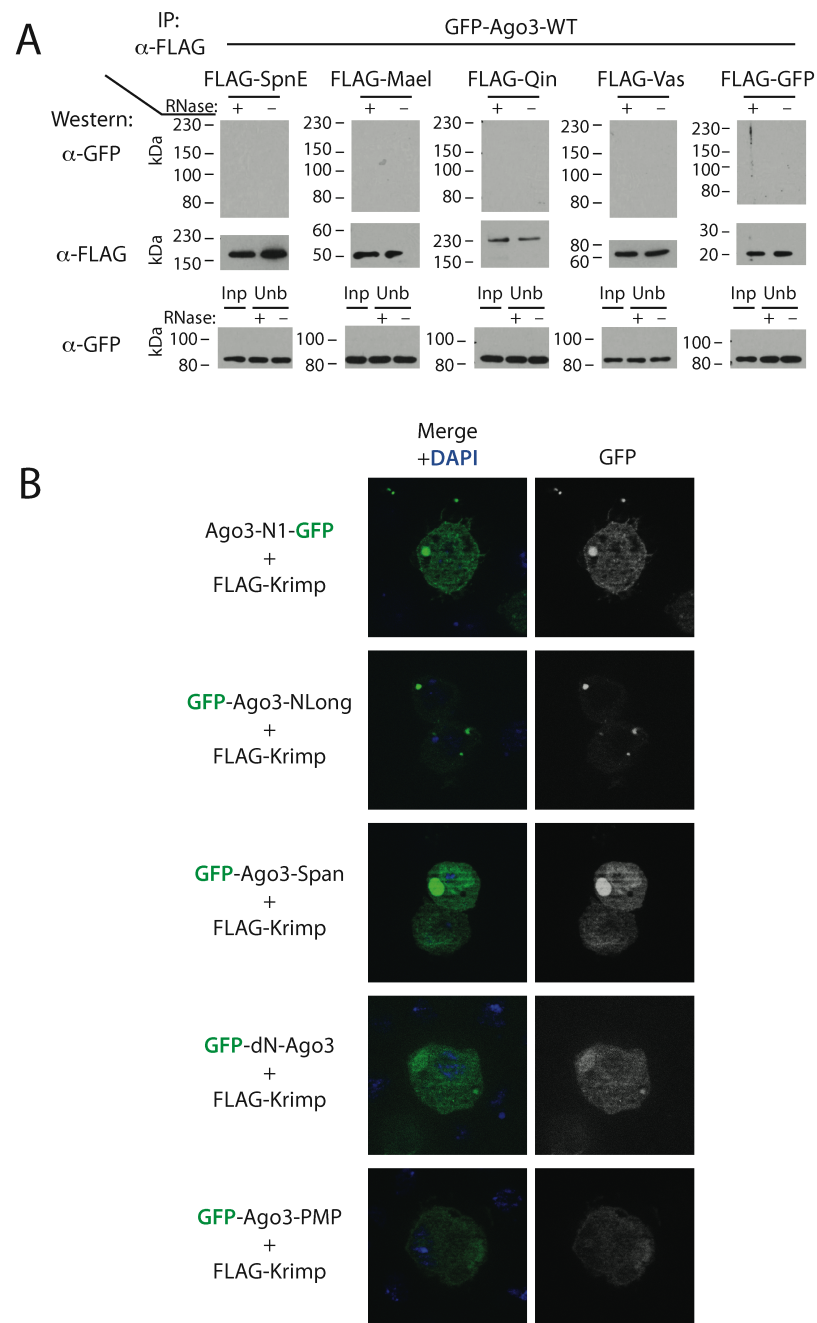
The N-terminal region of Krimp is required for the formation of granules GFP-tagged fragments of Krimp protein containing the first 310 amino acids of Krimp are able to granularize when co-expressed with FLAG-tagged full length Krimp (Krimp-FL), but fragments not containing these residues have diffused cytoplasmic localization in S2 cells. DNA is stained with DAPI (blue).

Webster et al. Supplemental Figure 3



***Supplemental Figure 4: Ago3 does not co-immunopurify with different nuage components in ovaries; Granularization of Ago3 fragments in S2 cells.***

(A) Ago3 does not interact with SpnE, Mael, Qin, or Vas proteins in *Drosophila* ovary. Immunoprecipitations were performed from ovaries of transgenic flies that express FLAG- and GFP-tagged proteins in germ cells under control of the MaG4 driver. FLAG-tagged proteins were immunoprecipitated in the presence (+) or absence (-) of RNase A followed by Western blotting to detect GFP- and FLAG-tagged proteins. The lower panel shows Input (Inp) and the unbound fractions (Unb) collected after immunoprecipitation. FLAG-GFP protein was used to confirm specificity of interactions, as it did not interact with Ago3. (B) The N-terminal region of Ago3 is required to be recruited to Krimp granules in S2 cells. The localization of GFP-tagged Ago3-NLong fragment is granular when co-expressed with Krimp, while all other fragments are uniformly distributed in the cytoplasm. S2 cells were transfected with plasmids encoding GFP-tagged Ago3 fragments (green) in combination with FLAG-tagged Krimp protein. DNA is stained with DAPI (blue).

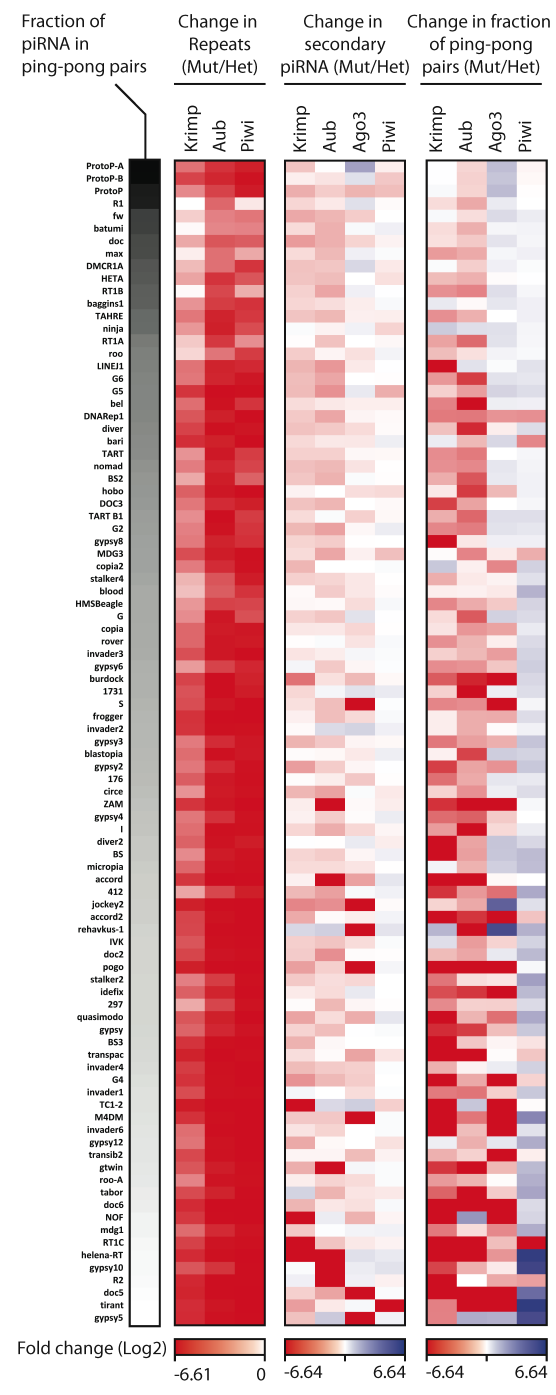




***Supplemental Figure 5: The effect of Krimp, Aub, Ago3 and Piwi mutations on ping-pong piRNA processing.***

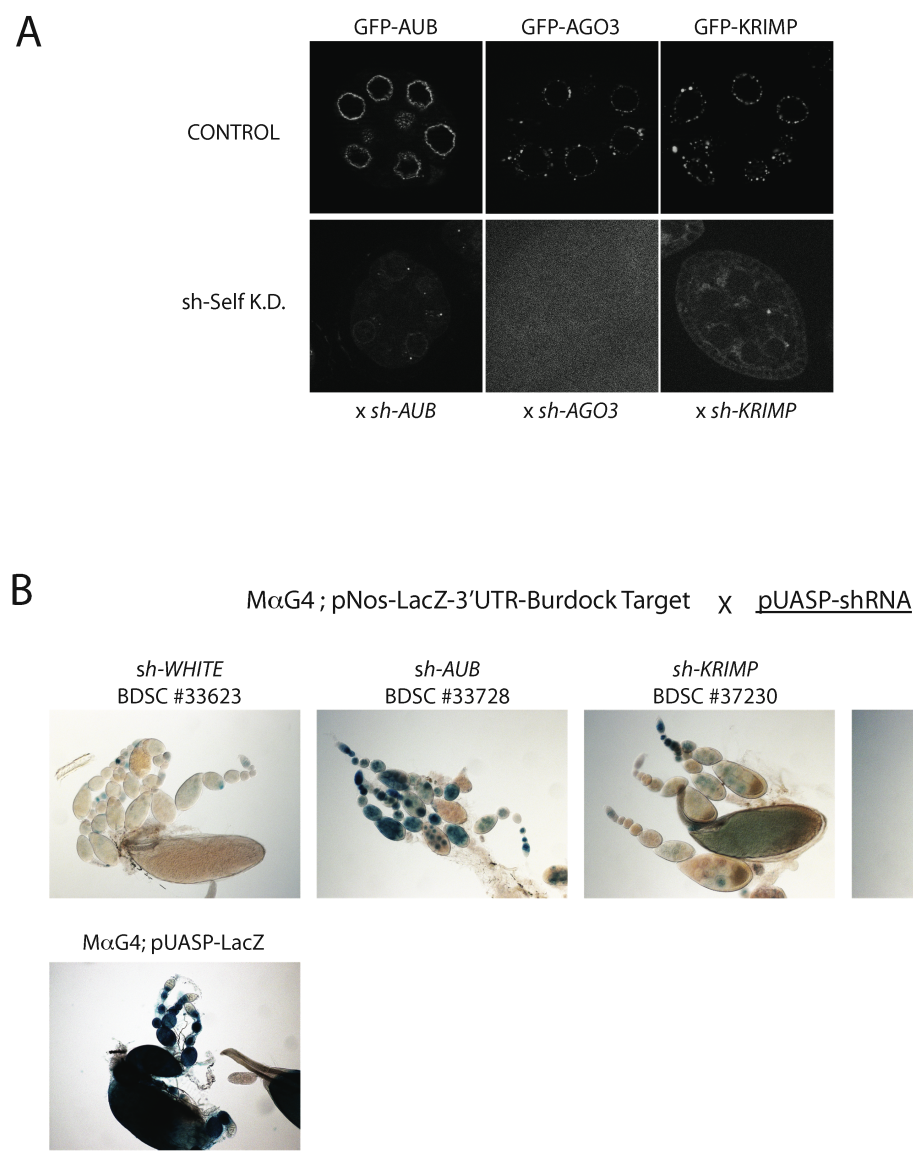
The fraction of secondary piRNAs and the fraction of the ping-pong pairs that have 10nt overlap were analyzed in Krimp, Piwi, Aub, and Ago3 mutants and control (heterozygous) flies. The changes in mutant compared to control for each family of transposable elements are shown on the heat map. TE families with at least 400 piRNA reads in control libraries are included.

Webster et al. Supplemental Figure 5



***Supplemental Figure 6. Evaluating the specificity and efficacy of knockdown fly strains.***

(A) Knockdown lines show specific knockdown of GFP-tagged proteins. Transgenic flies expressing GFP- tagged Aub, Ago3, or Krimp were crossed to shRNA lines targeting the respective GFP tagged protein. The GFP transgene and shRNA were co-expressed inside the ovary using the same driver. (B) Knockdown lines show effective knockdown by derepressing a LacZ sensor responsive to loss of Burdock piRNA. The Burdock sensor containing the MaG4 driver was crossed to knockdown lines: as a control, sh-White did not derepress LacZ, while knockdown of Aub, Ago3, and Krimp induced derepression of the LacZ sensor.



## CHAPTER 4

### **Arginine methylation signals the piRNA-loaded states of Aub and Ago3 to enable interactions with Tudor domain proteins**

This chapter will be published as an individual manuscript.

## INTRODUCTION

The *Drosophila* Piwi proteins, Piwi, Aubergine (Aub), and Argonaute-3 (Ago3), share similar protein architectures consisting of an N-terminal region, PAZ, middle (MID), and PIWI domain that have strong conservation of protein sequence. In spite of their similarities, they localize to different compartments of germ cells, and have different functions in the piRNA pathway and distinct fates upon the completion of oogenesis. While Piwi and Aub are loaded with primary piRNA enriched for sequences anti-sense to TE mRNA, Ago3 is the recipient of secondary piRNA from cleaved sense-strand TE transcripts. Aub and Ago3 are cytoplasmic proteins; however, Piwi localizes to the nucleus of germ cells. Its function in the nucleus is to target the deposition of transcriptionally repressive chromatin marks to genomic loci expressing active TE copies [101, 103]. Conversely, Aub and Ago3 cooperate to destroy TE transcripts in the piRNA ping-pong pathway in nuage. The fates of each protein are also unique: Aub accumulates at the posterior pole plasm of the oocyte at later stages of oogenesis, while Piwi persists within the syncytial cytoplasm of the oocyte without a specific localization pattern. Unlike Aub and Piwi, the levels of Ago3 protein are greatly reduced in the oocyte. The mechanism that allows the three *Drosophila* Piwi proteins to be differentiated by cellular machinery remains to be understood.

Unlike Argonaute-family proteins, Piwi-family proteins are targets for arginine methylation. The cytoplasmic PRMT5 methylosome complex composed of Capsul  n (Csl, dart5) and its co-factor, Valois (Vls, MEP50) modify arginines in arginine-rich motifs into symmetrically dimethylated arginines (sDMA)[68, 76]. Many piRNA pathway components have the potential to bind sDMA residues through interactions with Tudor domains, which are ubiquitously found in many protein factors of the piRNA pathway (see Chapter 1). While piRNA pathway-associated proteins contain Tudor domains, the role sDMA modifications play in regulating Piwi protein function remains elusive. sDMA modifications were identified within the N-terminal regions of Piwi, Aub, and Ago3 [109, 123]. Indeed, several cognate Tudor domains for methylated Aub and Ago3 have been identified, including Papi and the eponymous Tudor protein (Tud), which interact with methylated Ago3 and Aub, respectively [69, 70, 72, 75, 109].

However, the general function of arginine methylation as a regulator of Piwi protein function is not yet clearly defined. Experiments in a cultured cell model for the silk-worm germline revealed that N-terminal arginine methylation motifs do not influence the loading of piRNA into Ago3 [124]. Instead, piRNA loading appears to be dictated by the Mid-Piwi domains of Piwi proteins[125].

In Chapter 3 of this thesis, I present evidence that the underlying interaction of the piRNA ping-pong pathway between Aubergine (Aub) and Argonaute-3 (Ago3) is bridged by Krimper protein. The N-terminal regions of Aub and Ago3 are able to bind the two Tudor domains of Krimper, while Krimper forms dimeric or multimeric complexes with itself to simultaneously accommodate Aub and Ago3 proteins. In addition to the formation of an Aub-Ago3 complex, the function of the ping-pong mechanism requires that one Piwi protein be loaded with piRNA so as to target complementary transcripts and the other to be unloaded and primed for loading of cleaved target transcripts. As our model suggests, the arrangement of Piwi proteins in the Krimper complex are able to position Aub protein loaded with piRNA adjacent to unloaded Ago3 protein. However, it remains to be understood how Krimper differentiates Aub and Ago3, both existing in piRNA-loaded or unloaded states.

In this chapter, we present evidence confirming that the N-terminal region and N-terminal arginine methylation of *Drosophila* Piwi proteins does not dictate the loading of piRNA into Piwi proteins. Instead, we propose that methylation of N-terminal arginine residues signals that a Piwi protein is loaded with piRNA, enabling downstream interactions with specific Tudor-domain proteins. We show that Aub methylation is not required for normal loading of piRNA, but is essential for piRNA pathway function and fertility. Specifically, methylation-deficient Aub is unable to interact with Krimper protein, while Ago3 interacts with Krimper independently of its methylation state. The role of Ago3 arginine methylation also appears to affect at least two piRNA pathway-associated proteins, Vasa and Krimper, which accumulate as more massive modified isoforms in ovaries expressing methylation-deficient Ago3 protein. Overall, we propose a comprehensive model where arginine methylation functions as a signal that is used by germ cells to differentiate loaded from unloaded Piwi proteins.

## RESULTS

### ***The N-terminal region is the most diverse region of Piwi proteins but is not essential for piRNA loading***

Piwi proteins contain defined PIWI, PAZ and Mid domains responsible for target cleavage, and binding of the 3' and 5' end of piRNA, respectively [93, 94, 96, 118]. Upstream of the PAZ domain is the N-terminal region of the proteins of which the structure is not known. The sequence of the N-terminal region is the most variable domain between Piwi members (Figure 1A). Furthermore, the ratio of non-synonymous (dN) to synonymous (dS) mutation rates is moderately elevated across the N-terminal region of Aub and Ago3, indicative of possible positive selection (Figure 1B, 1C) [126-129]. To investigate the role of the N-terminal domain in the localization and in piRNA loading of Piwi proteins, we generated a series of domain swaps that reciprocally exchanged the first 100 amino acids of Aubergine with other Piwi proteins. Each chimeric protein was N-terminally tagged with GFP to visualize their localization in addition to using GFP as an immunopurification tag. We first tested the contribution that the N-terminal regions of Piwi, Aub, and Ago3 have on localization. Piwi localizes to the nucleus of ovarian nurse cells. Grafting the N-terminal region of Piwi onto the PAZ-mid-PIWI body of Aub (N-Piwi-Aub) resulted in recruitment of the chimeric protein to the nucleus. However, the reciprocal chimeric protein consisting of the N-terminal region of Aub grafted onto the body of Piwi (N-Aub-Piwi) did not localize to the nucleus, but instead localized to the perinuclear nuage compartment (Figure 2A). This finding is in agreement with a previous study that reported that an N-terminally truncated Piwi mutant was unable to localize to the nucleus, but instead was diffused in the cytoplasm of germ cells [130]. Therefore, the Piwi N-terminal region functions independently of the PAZ-mid-PIWI body as a nuclear localization signal.

While N-terminally truncated Piwi localizes to the cytoplasm, the addition of the Aub N-terminal region was able to recruit it to nuage similarly to wildtype protein (Figure 2A). However, when the N-terminal region of Aub was expressed alone, it localized throughout the cytoplasm and nucleus but did not have a clear localization pattern to nuage. The PAZ-mid-PIWI body of Aub also appears to contribute to Aub



nuage localization, since N-terminally truncated Aub (dN-Aub) protein localized in part to nuage, but also diffused throughout the cytoplasm (Figure 2B). We also generated a chimeric protein consisting of the N-terminal region of Ago3 grafted onto the body of Aub (N-Ago3-Aub) and found that it localized similarly to either Aub and Ago3 (Figure 2A). Expression of a reciprocal chimeric protein consisting of the Aub N-terminal region fused to Ago3 was not detected by microscopy. Ago3 N-terminal truncation (dN-Ago3) did not appear to express either. It is possible that the N-terminal region of Ago3 is important for protein stability.

To test if the N-terminal regions of Piwi proteins contribute to piRNA loading, we purified small RNA from immunoprecipitations of the domain-swapped proteins. Indeed, all proteins appeared to bind piRNA of sizes 24-29nt in length (Figure 2C). The amount of associated piRNA was estimated using a semi-quantitative approach that compared protein levels by western blot to radiolabeled RNA, utilizing a synthetic 42M oligomer to control for differences during RNA purification (see materials & methods). The quantity of associated piRNA was decreased in domain-swapped proteins, which is likely a result of perturbing the native protein structures (Figure 2D). Interestingly, the levels of associated piRNA of N-terminal truncated Aub (dN-Aub) were not greatly reduced.

To determine the contribution of N-terminal regions on piRNA loading, we analyzed the sequences of associated piRNA. piRNA are generally between 24 – 30nt in length, although each Piwi protein are loaded with piRNA of distinct median lengths [99]. Piwi-associated piRNA tend to be ~26nt in length, while the median length of Aub-associated piRNA is ~25nt and Ago3, ~24nt in length. Analysis of piRNA associated with N-terminal region domain swaps reveals that piRNA bound to N-Aub-Piwi is most similar to Piwi-bound piRNA, while N-Piwi-Aub is more similar to Aub-bound piRNA in length (Figure 2E). Overall, the median length of piRNA is dictated by the identity of the PAZ-mid-PIWI body of the protein, and not by the N-terminal region.

We examined if N-terminal regions contributed to the strand orientation and nucleotide biases of associated piRNA. Both Aub and Piwi load with piRNA that is predominantly anti-sense to annotated TE sequences and have a preference for uridine at the 5' nucleotide (1U bias), while most Ago3-associated piRNA is of sense orientation

with a bias for adenine at position 10 (10A bias) [38, 99]. The strand orientation of piRNA bound to N-terminal domain-swapped proteins was most similar to the protein coding the PAZ-mid-PIWI body compared to the identity of the N-terminal region. For example, piRNA bound to N-Ago3-Aub were mostly of anti-sense orientation characteristic of Aub-bound piRNA (Figure 2F). The nucleotide biases of the domain-swapped proteins also followed this observation, where piRNA bound to N-Ago3-Aub had a strong 1U bias but not a 10A bias, which is typical of Aub, but not Ago3-bound piRNA (Figure 2G). In summary, our findings show that the N-terminal regions of Piwi, Aub, and Ago3 have a minor role in piRNA loading.

To exclude the possibility that differences in localization and piRNA loading are masked by domain swaps between related Piwi proteins that operate in the same pathway we tested N-terminal domain swaps with Aub and Argonaute-1 (Ago1) to generate two chimeric proteins, N-Aub-Ago1 and N-Ago1-Aub. Ago1 is the effector protein for the *Drosophila* microRNA (miRNA) pathway and *Drosophila* miRNA are 22 nt in length. A previous study reported that N-terminal residues of protozoan Ago1 are essential for function, although it is unknown if the N-terminal region of *Drosophila* Ago1 would impart loading of miRNA into the grafted body of Aub, or whether the N-terminal region of Aub could load piRNA into Ago1. Since *Drosophila* have two isoforms of Ago1—Ago1Ra and Ago1Rb—that differ in the length and composition of their N-terminal regions, we generated domain swaps with Aub for both regions (N-Ago1Ra-Aub and N-Ago1Rb-Aub) (Figure 1B). The localization of wild type Ago1 proteins was most similar to N-Aub-Ago1 and was present throughout the cytoplasm of nurse cells, but also localized to the perinuclear space. Conversely, both N-Ago1Ra-Aub and N-Ago1Rb-Aub appeared most concentrated in perinuclear nuage, similar to wildtype Aub protein (Figure 3A).

The N-terminal region of chimeric proteins did not impact the length of small RNA associated with each protein, since N-Aub-Ago1 purified with 22 nt RNA, even though it contained the N-terminal region of Aub. Similarly, N-Ago1Ra-Aub and N-Ago1Rb-Aub purified with small RNA identical in size distribution to Aub piRNA (Figure 2C, 3B). Chimeric proteins between Aub and Ago1 also co-purified with less small RNA than wildtype proteins, with the exception of N-Ago1Ra-Aub, which purified

with more than twice as much small RNA compared to wildtype Aub (Figure 2D). Through mapping of cloned small RNA purified from each chimeric protein, we found that chimeric proteins containing the PAZ-mid-PIWI of Aub were bound to piRNA and not miRNA, while N-Aub-Ago1 purified with an abundant population of miRNA (Figure 3C). We compared the populations of miRNA purified from Ago1-containing chimeric proteins to further understand the differences between the two isoforms of wild type Ago1 and the chimeric N-Aub-Ago1. Overall, any differences appeared to be independent of miRNA abundance, although Ago1Ra preferentially loaded with several miRNA, including bantam, mir-308, -304, -965, and -999, compared to Ago1Rb, which had a higher abundance of mir-2a-2 compared to Ago1Ra. The chimeric N-Aub-Ago1 had lower levels of mir-9c, -33, -306, and 31b, but elevated levels of mir-13b-2 compared to either Ago1Ra or Ago1Rb (Figure 3D). We did not find any correlation between the levels of these miRNA and any reported functions in germ cells.

In conclusion, our data clearly establishes that piRNA loading is not influenced by the composition of N-terminal regions. Instead, the specificity of piRNA length, strand orientation, and nucleotide bias are imparted by the identity of PAZ-mid-PIWI modules. In only some instances, the N-terminal region of Piwi proteins regulate protein localization. The Piwi N-terminal region functions as a nuclear localization signal even when grafted to the body of Aub protein that is loaded with Aub-specific piRNA. In contrast, the localization of Aub protein is independent of the N-terminal region, although the N-terminal region might enhance localization to perinuclear nuage.

### ***N-terminal arginine methylation is not required for localization or piRNA loading of Aub or Ago3***

The N-terminal regions of Aub and Ago3 contain several RG/RA repeat-motifs that are targets of symmetrical dimethyl-arginine (sDMA) modifications [69, 109, 123]. Methylated arginine residues of Aub provide a binding interface for interaction with the Tudor domain of Tud protein [69, 70, 75, 131, 132]. Of the numerous proteins that were genetically implicated to be factors of the piRNA pathway and that localize to nuage, many have one or several Tudor domains [86]. It was proposed that interactions between modified Piwi proteins and Tudor domain proteins could be responsible for nuage assembly [133-135].

To test the role of arginine methylation in the localization and function of Piwi proteins, we generated mutant versions of GFP-tagged Aub and Ago3 by replacing N-terminal arginine residues in RG/RA repeat-motifs with lysine residues (Aub-RK and Ago3-RK) (Figure 1B, 1C). The subcellular localization of Ago3-RK is indistinguishable from wild-type Ago3, indicating that arginine methylation is dispensable for its localization in nuage (Figure 4A). Aub-RK is also localized to nuage; however, a deficiency in arginine methylation eliminated its accumulation at the pole plasm (Figure 4A; 4B). Aub localization to the pole plasm was reported to be dependent on Tudor protein that interacts with sDMA modifications on Aub [70, 75, 109, 131, 132]. Indeed, while wild-type Aub co-immunoprecipitates with Tudor from ovary extracts, Aub-RK does not interact with Tud (Figure 4C). Arginine methylation seems to destabilize Ago3, since Ago3-RK protein was approximately seven-fold more abundant than wild-type protein when corresponding transgenes inserted in the same genomic position were expressed using the same GAL4 driver (Figure 4D). This was not the case for Aub-RK, which showed approximately 30% lower levels to wild-type Aub.

In contrast to a strong requirement for arginine methylation in the localization of Aub to the pole plasm, arginine methylation has only minor effects on Aub localization to nuage: overall the nuage localization of Aub-RK is preserved, although we noticed a more diffuse localization in the cytoplasm compared to wild-type protein (Figure 4A). The analysis of the ratio of nuage to cytoplasm fluorescence intensity showed that Aub-RK was less enriched in nuage compared to the wild-type protein (Data not shown). A similar increase in diffuse cytoplasmic localization was observed for wild-type Aub when expressed in mutants lacking the PRMT5 methylome components, Capsul  n (*Csul*; *PRMT5*) and Valois (*Vls*; *MEP50*), both of which are required for sDMA modifications (Figure 5A, 5B) [68, 69, 76, 136]. Using FRAP we determined that the fraction of mobile Aub-RK protein is approximately 10% higher than that of wild-type protein (Figure 4E). Overall, we conclude that arginine methylation has a minor influence on the localization and dynamics of Aub and Ago3 in nuage granules, while it is critical for recruitment of Aub to the pole plasm through interactions with Tudor protein.

Next we analyzed the role of Aub and Ago3 arginine methylation in the loading of piRNAs into either protein. The levels of piRNAs bound by purified proteins were

similar in methylation-impaired Aub and Ago3 proteins compared to their wild-type counterparts (Figure 4F; 4G). We sequenced piRNAs associated with Aub-RK and found that the piRNA profile was similar to wild-type protein, indicating that Arginine methylation is indeed dispensable for piRNA loading of both Aub and Ago3 (Figure 4H). Strand orientation and nucleotide biases typical of Aub protein were unchanged for Aub-RK (Figure 4I). Overall, arginine methylation does not affect piRNA loading of either Aub or Ago3.

In summary, we show that the N-terminal regions of Aub and Ago3 that harbor arginine methylation sites recognized by Tudor domain proteins are largely dispensable for both nuage localization and piRNA loading. Despite the N-terminal region being the least similar in sequence between Aub and Ago3, it is not responsible for differentiating the roles of Aub and Ago3 in the piRNA pathway.

***Aub arginine methylation is essential for progression of the piRNA ping-pong pathway and fertility***

Since arginine methylation does not impact the loading of piRNA into Aub, we wondered if it is essential for the function of the piRNA pathway. To further investigate the role of Aub arginine methylation, we generated a GFP-tagged Aub-RK transgene under its native regulatory elements. GFP-tagged Aub expressed under control of native regulatory elements shows expression from region one of the germarium onward to the late stage oocyte and rescues the *aub* mutant phenotype. Furthermore, its localization in nuage and in the pole plasm was identical to GFP-tagged Aub expressed by the Maternal  $\alpha$ -Tubulin 67C ( $M\alpha G4$ ) driver (Figure 6A)[112]. We successfully rescued the Aub mutant background using GFP-tagged Aub-WT under the expression control of native regulatory elements, so we tested if GFP-Aub-RK was capable of rescuing the Aub mutant background. Surprisingly, GFP-Aub-RK was not able to rescue the Aub mutant background, as flies of this genotype are sterile. Early stages of Aub-RK rescue ovaries appear to have intact egg chambers, although late stage chambers experience fusions of two egg chambers or gross morphological defects (Figure 6B). This indicated that while methylation is not necessary for Aub nuage localization or piRNA loading, it is essential for the TE repression function of the piRNA pathway.

In Chapter 3, we present how the piRNA ping-pong complex assembles. In this complex, Aub and Ago3 interactions are bridged by the Tudor domain-containing protein, Krimper (Krimp). The data suggested a model where Aub protein loaded with piRNA would select RNA targets to be loaded into Ago3, since Krimp appeared to only interact with Ago3 that was devoid of piRNA. It remained unclear how Krimp distinguishes Aub and Ago3, with each protein existing in their respective piRNA-loaded and unloaded states. Molecular dissection revealed that Aub and Ago3 both interact with the Tudor domains of Krimp through their N-terminal regions. Since the Aub-RK protein was unable to rescue the Aub mutant phenotype, we tested the impact of Aub arginine methylation on the stability of the Krimp-Ago3 complex. Indeed, the Krimp-Ago3 complex appears to be stabilized when Aub arginine methylation is disabled, suggesting that Aub methylation is required to bind Krimper and this interaction is essential for progression of the ping-pong pathway (Figure 6C). To verify if Aub or Ago3 interactions with Krimp are dependent on arginine methylation, we expressed Aub-RK or Ago3-RK tagged with GFP in S2 cells, which do not express most piRNA pathway components but express PRMT5 methylosome complex components. While wildtype version of Aub and Ago3 granulate in the presence of Krimp protein in S2 cells, we noticed that methylation deficient Aub no longer localized to granules. Conversely, arginine-methylation deficient Ago3 localized to granules much like wildtype protein (Figure 6D). Co-immunoprecipitation experiments confirmed that while both wildtype and methylation deficient Ago3 co-purified with Krimp, methylation deficient Aub was unable to interact with Krimp (Figure 6E). While arginine methylation is not essential for localization, or piRNA loading, it is required for Aub to interact with the Krimp-Ago3 complex during ping-pong, as this interaction is essential for fertility.

***piRNA loading is required for arginine methylation***

Since N-terminal arginine methylation is dispensable for the loading of Aub and Ago3 but is required for the function of the piRNA pathway, we sought to understand the requirements for arginine methylation Aub and Ago3 methylation. Arginine methylation is either an unregulated modification that is added to Aub and Ago3 constitutively, or it is a regulated process that requires Aub and Ago3 to undergo a conformational change through loading with piRNA, catalytic cleavage of target RNA, or another unknown

process. We examined the methylation states of mutant proteins of Aub and Ago3 that could not bind piRNA (see Chapter 3) or were deficient in endonucleolytic function (see Chapter 5). Reagents that detect methylated Aub and Ago3 have poor specificity and were unable to confirm the methylation states of the mutant proteins. Instead, we examined Aub localization at the pole plasm, which requires arginine methylation to interact with Tud protein (Figure 4B, 4C). Interestingly, while the pole plasm localization of wild-type and catalytically inactive Aub were similar, piRNA binding-deficient Aub (Aub-YK) was unable to localize to the pole plasm, showing the same defect as methylation-deficient Aub-RK (Figure 7A). This indicated that while catalytic activity is dispensable for Aub localization in the oocyte, Aub requires loading with piRNA to be methylated and to interact with Tud protein at the pole plasm.

Since Ago3 does not localize to the pole plasm, we used a different assay to test the methylation state of Ago3. We found that arginine methylation of Ago3 is necessary for normal oogenesis. piRNA pathway mutants phenotypically exhibit increased Chk2 kinase activity and cell-cycle arrest in response to transposon-induced DNA lesions. Another consequence of Chk2 activation is post-translational modification of Vasa protein [107, 137]. Placing a pathway mutant on a Chk2 mutant background can attenuate cell cycle arrest signals and ‘rescue’ the modified Vasa isoform [138]. Interestingly, expressing Ago3-RK over a wild type background induced post-translational modification of Vasa but did not observe accumulation of the modified Vasa isoform for Ago3-WT or Ago3-YK, nor any of the Aub transgenic proteins (Figure 7B). Ago-RK and Ago3-CD also appeared to induce the modification of Krimp protein, which was previously never been reported (Figure 7C). The accumulation of modified Vasa protein did not appear to alter Vas localization to nuage or the pole plasm, confirming results from another study (Figure 7D)[138].

Stable lines of MaG4; GFP-Ago3-RK, and MaG4; GFP-Ago3-CD are not homozygous for the transgenic AGO3 insertion, which could indicate that Ago3-RK and Ago3-CD are dominant negative alleles that induce a DNA damage signal in the developing oocyte. We tested if modified Vasa protein could be ‘rescued’ or reduced in levels by reducing the levels of Chk2 by crossing flies expressing Ago3-RK and Ago3-CD to an *sh-mnk* knockdown line, but did not observe a decrease in the modified Vasa

isoform (Figure 7E). This could be a consequence of the *sh-mnk* line being ineffective, although the resulting knockdown progeny were sterile. We tested if the modified Vasa isoform was a consequence of phosphorylation, so we treated lysate with calf intestinal phosphatase to check for a reduction in the modified Vasa isoform. Ago3-CD responded to phosphatase treatment, but the levels of modified Vasa isoform in flies expressing Ago3-RK did not appear to change (Figure 7F). Overall, our data show that Ago3 dominantly regulates the accumulation of modified Vas protein, while other piRNA pathway proteins are recessive regulators. Specifically, accumulation of modified Vas protein is responsive to Ago3 arginine methylation and its endonucleolytic function. piRNA loading of Ago3 does not cause accumulation of the modified Vas isoform, which implies Ago3 methylation occurs downstream of loading.

In summary, both Aub and Ago3 require piRNA loading before they can be methylated. Aub that is unable to bind piRNA does not accumulate at the pole plasm, which is an interaction that requires methylation of N-terminal arginine residues. Similarly, while methylation deficient Ago3 induces the accumulation of modified isoforms of Vas and Krimp, piRNA binding deficient Ago3 does not.

## DISCUSSION

### ***Arginine methylation of Aub and Ago3 N-terminal residues is largely dispensable for their recruitment to nuage and for loading with piRNAs***

Among the different members of Piwi-clade Argonaute proteins, the most diverse region is the N-terminal fragment upstream of the PAZ domain (Figure 1A). While no function was assigned to the N-terminal region in members of the Argonaute-clade, many members of the Piwi-clade in both invertebrates and vertebrates species harbor RG/RA motifs in their N-termini [123, 139]. These motifs were shown to be targets for symmetric arginine methylation by the PRMT5 methylosome complex [140, 141]. Methylated arginine residues in Piwi proteins provide an interface for binding Tudor domains [69, 70, 109, 133]. Several proteins with Tudor domains were genetically implicated in the piRNA pathway and were also shown to localize to nuage and associate with Piwi proteins [42, 70, 86, 142-144]. In fact, interactions between Piwi proteins and



Tudor domain proteins mediated by arginine methylation were proposed to be the main mechanism that allows assembly of multiple components into nuage granules [133-135].

We tested the impact of the N-terminal region on localization and piRNA loading of Piwi proteins. piRNA loading of Aub and Piwi depended not on the identity of the N-terminal region, but instead on the PAZ-mid-PIWI domains. This observation also held true for Ago1, which was loaded with miRNA even when the Aub N-terminal region was exchanged with its native N-terminal region (Figure 2C, 2E, 2F, 2G, 3B, 3C). In Chapter 3, we identified that the Ago3 interacts with Krimp through its N-terminus. It was therefore a surprise that the N-terminal region of Ago3 also did not influence the strand orientation or nucleotide bias of piRNA associated with the N-Ago3-Aub chimera.

Our results show that arginine methylation of Aub and Ago3 proteins plays a minor role in their recruitment and retention in nuage. Mutation of arginine residues that are targets for methylation in Aub results in a loss of Aub – Tud interactions, confirming that the interaction between Tud and Aub depends on an arginine methylation motif in its N-terminal region (Figure 4C). Furthermore, deficiency in Aub methylation impaired its accumulation at the pole plasm of the developing oocyte, supporting previous findings that methylation-dependent Tudor binding is required for pole plasm recruitment of Aub (Figure 4B) [69, 74, 109].

While Aub arginine methylation is required for its accumulation at the pole plasm, our results show that arginine methylation has only a minor role in Aub localization to nuage: overall, the nuage localization of Aub-RK mutant was not disrupted. However, we observed that a larger fraction of Aub-RK was diffuse in a cytoplasmic pool of protein and in nuage, a higher amount of protein was mobile. Considering the fact that Aub strongly co-localizes with Tud in nuage, this interaction plays at least some role in retaining Aub in nuage and reduces its exchange with the diffuse cytoplasmic fraction of protein. Nonetheless, we find that this is not the main mechanism that recruits Aub to nuage. As discussed in Chapter 3, a piRNA guide is required for Aub nuage localization. We show that the PAZ-mid-PIWI domains of Aub are sufficient to for piRNA loading, and as a result N-terminally truncated Aub localizes to nuage similarly to wild type Aub.

Arginine methylation seems to play an even less significant role in Ago3 localization to nuage, as we were not able to find differences between the steady-state localization or mobility of wild type and arginine-methylation mutant Ago3 proteins. Importantly, the absence of strong defects in arginine methylation mutant proteins is not caused by our failure to mutate all potential sites of arginine methylation, as we observed similar Aub and Ago3 localization in ovaries of flies mutant for PRMT5 methylome components, Csu1 and Vls, which abolish the function of the pathway responsible for cytoplasmic sDMA modifications [69]. Our study reveals one important function of the N-terminal sequence in Ago3: the protein stability of Ago3, but not Aub, is influenced by arginine methylation, as we observed the stabilization of mutant Ago3-RK protein in ovaries. Furthermore, N-terminally truncated Ago3 protein could not be detected when expressed in germ cells, supporting the role of the N-terminal region in regulating protein stability.

Arginine methylation does not play a role in the loading of Aub with piRNAs, as the quantity and profile of piRNAs loaded into protein seem to be identical between methylation-site mutant and wild type proteins (Figure 4F; 4G; 4H). Furthermore, complete deletion of the N-terminal fragment in Aub or its replacement with N-terminal fragment of Ago3 does not affect protein localization and loading with piRNAs (Figure 2B, 2C, 2F, 2G, 4H; 4I). Overall our results indicate that the distinct roles of Aub and Ago3 proteins in the piRNA pathway are likely mediated by the C-terminal portions of the proteins, which include the conserved PAZ, Mid, and PIWI domains. The negligible role of the N-terminus seems to be a conserved feature of Piwi proteins, as a similar conclusion was recently reached by investigating silkworm Piwi proteins [125].

### ***Arginine methylation occurs downstream of piRNA loading***

Based on our data, we reason that arginine methylation is a signal that indicates when Aub and Ago3 are loaded with piRNA. We show that Aub loads with piRNA independently of arginine methylation. However, Aub methylation is required for interactions with Krimp protein, which is the molecular bridge that recruits unloaded Ago3 during ping-pong (Figure 6D, 6E). The essential role of Aub methylation is exemplified by the inability of Aub methylation mutant to rescue fertility of Aub mutant ovaries (Figure 6B). Interestingly, the defects observed in the Aub-RK rescue are less

severe than Aub mutant flies lacking the Aub-RK protein, which have ovaries that are almost completely atrophied. This might indicate that the Aub-RK transgene does impart some protection against TE in early stages of oogenesis. However, because methylation deficient Aub cannot interact with Krimp protein to direct Ago3 loading during ping-pong, there is a lack of piRNA amplification for sequences targeting active TE. The resulting population of piRNA in Aub-RK rescue ovaries is insufficient for immunity against TE.

Nevertheless, while methylation-deficient Aub is able to load with piRNA, downstream arginine methylation is required for critical interactions with the Tudor domains of Krimp and without methylation, Aub cannot participate in critical interactions with Tudor domain containing proteins. There would not be any biological reason for unloaded Aub to interact with Krimp, and it is more likely that Aub is methylated only after loading. Another reason why methylation of Aub is a regulatory step downstream of piRNA loading is that since piRNA binding-deficient Aub does not localize to the pole plasm, an interaction dependent on arginine methylation (Figure 7A).

Similar to Aub, loading of piRNA into Ago3 does not require methylation and its interaction with Krimp— a factor that contributes to Ago3 loading— is also independent of methylation. However, Ago3 arginine methylation is essential for normal behavior of piRNA pathway components. When the methylation site of Ago3 is mutated, two essential piRNA pathway components, Vasa and Krimper, are modified into more massive isoforms (Figure 7B, 7C). Unlike Ago3-RK, piRNA binding deficient Ago3 does not elicit accumulation of Vas or Krimp modified isoforms. This suggests that loading of Ago3 with piRNA precedes methylation. Since Ago3 lacking catalytic residues responsible for endonucleolytic function also causes accumulation of modified Vas and Krimp isoforms, arginine methylation appears to be closely related, or dependent on Ago3 endonucleolytic function. The biological reason for why Ago3 influences modification of Vas and Krimp remains unknown. Furthermore, the function of modified isoforms of either Vas and Krimp are also poorly understood. It remains to be understood how Ago3 methylation and endonucleolytic activity regulate the function of these piRNA pathway factors.

We envision that the sequential nature of Piwi protein function in the piRNA pathway could hypothetically operate under the simple principle that once a Piwi protein is loaded with piRNA it is then methylated (Figure 8A). Structurally, it is possible that loading of a Piwi with a piRNA guide could induce a conformational change in the protein to make N-terminal arginine residues accessible for methylation. If both Aub and Ago3 operated under this same principle, then we propose a model for the ping-pong pathway that operates as follows: (1) methylated Aub loaded with piRNA in a triple complex with unloaded, unmethylated Ago3 coordinated by Krimp scans transcripts for complementary to its guide piRNA. When a complementary transcript is targeted and cleaved, it is loaded into Ago3. (2) Ago3 loaded with piRNA dissociates from Krimp, and (3) is methylated due to a conformational change that makes N-terminal arginine residues accessible. (4) Methylated and piRNA-loaded Ago3 interacts with designated Tudor-domain containing proteins. One such protein is Papi, which contains a single Tudor domain and two KH-I domains which are known to interact with A-U rich containing RNAs. Transcripts originating from piRNA clusters are degenerate sequences with high A-U content. While the molecular biology of Papi is not yet understood, it is possible that it couples methylated, loaded Ago3 with unloaded, unmethylated Aub protein to direct loading of Aub. Indeed, Aub protein no longer localizes to nuage in Papi knockdown ovaries (data not shown). (5) Ago3 targets and cleaves a complementary transcript that is loaded into Aub. (6) Aub dissociates from this complex and is then methylated. The cycle is perpetuated throughout oogenesis. In the final steps of oogenesis, there are also two separate fates for methylated Aub and Ago3. We show that Aub methylation is required for accumulation at the pole plasm, where Aub can be inherited by the primordial germ cells of the embryo. Conversely, Ago3 methylation destabilizes the protein, as it serves no biological function in targeting TE being composed of sequences in sense orientation to TE transcripts. In addition, the ping-pong pathway has already produced a population of piRNA that are highly specified for only those sequences targeting active TE. Overall, this system theoretically works under the rules of a simple principle: that piRNA loading enables arginine methylation of Aub or Ago3 (Figure 8B). It will be interesting to see if this hypothetical model has credence, in addition to understanding if Piwi functions under the same principle.

## REFERENCES

1. Slotkin, R. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome*. Nature reviews. Genetics, 2007. **8**(4): p. 272-285.
2. Siomi, M., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nature reviews. Molecular cell biology, 2011. **12**(4): p. 246-258.
3. Carmell, M., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis*. Genes & development, 2002. **16**(21): p. 2733-2742.
4. Cox, D., et al., *A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal*. Genes & development, 1998. **12**(23): p. 3715-3727.
5. Kuramochi-Miyagawa, S., et al., *Mili, a mammalian member of piwi family gene, is essential for spermatogenesis*. Development (Cambridge, England), 2004. **131**(4): p. 839-849.
6. Kuramochi-Miyagawa, S., et al., *Two mouse piwi-related genes: miwi and mili*. Mechanisms of development, 2001. **108**(1-2): p. 121-133.
7. Lin, H. and A. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development (Cambridge, England), 1997. **124**(12): p. 2463-2476.
8. Houwing, S., et al., *A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish*. Cell, 2007. **129**(1): p. 69-82.
9. Song, J.-J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science (New York, N.Y.), 2004. **305**(5689): p. 1434-1437.
10. Wang, Y., et al., *Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex*. Nature, 2008. **456**(7224): p. 921-926.
11. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nature cell biology, 2009. **11**(5): p. 652-658.
12. Liu, K., et al., *Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(43): p. 18398-18403.

13. Nishida, K., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. The EMBO journal, 2009. **28**(24): p. 3820-3831.
14. Vagin, V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes & development, 2009. **23**(15): p. 1749-1762.
15. Huang, H.-Y., et al., *Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish*. The EMBO journal, 2011. **30**(16): p. 3298-3308.
16. Mathioudakis, N., et al., *The multiple Tudor domain-containing protein TDRD1 is a molecular scaffold for mouse Piwi proteins and piRNA biogenesis factors*. RNA (New York, N.Y.), 2012. **18**(11): p. 2056-2072.
17. Aravin, A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Molecular cell, 2008. **31**(6): p. 785-799.
18. Aravin, A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS genetics, 2009. **5**(12).
19. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-1103.
20. Malone, C.D., et al., *Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary*. Cell, 2009. **137**(3): p. 522-35.
21. Aravin, A., et al., *A novel class of small RNAs bind to MILI protein in mouse testes*. Nature, 2006. **442**(7099): p. 203-207.
22. Girard, A., et al., *A germline-specific class of small RNAs binds mammalian Piwi proteins*. Nature, 2006. **442**(7099): p. 199-202.
23. Grivna, S., et al., *A novel class of small RNAs in mouse spermatogenic cells*. Genes & development, 2006. **20**(13): p. 1709-1714.
24. Lau, N., et al., *Characterization of the piRNA complex from rat testes*. Science (New York, N.Y.), 2006. **313**(5785): p. 363-367.
25. Aravin, A.A., et al., *Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control*. Science, 2007. **316**(5825): p. 744-747.
26. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes & development, 2013. **27**(4): p. 390-399.

27. Sienski, G., D. Dönertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-980.
28. Ro, S., et al., *Cloning and expression profiling of testis-expressed piRNA-like RNAs*. RNA (New York, N.Y.), 2007. **13**(10): p. 1693-1702.
29. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Molecular cell, 2013. **50**(5): p. 749-761.
30. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Molecular cell, 2013. **50**(5): p. 762-777.
31. Muerdter, F., et al., *A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila*. Molecular cell, 2013. **50**(5): p. 736-748.
32. Olivieri, D., et al., *An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila*. The EMBO journal, 2010. **29**(19): p. 3301-3317.
33. Ipsaro, J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-283.
34. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-287.
35. Voigt, F., et al., *Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc*. RNA (New York, N.Y.), 2012. **18**(12): p. 2128-2134.
36. Szakmary, A., et al., *The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster*. The Journal of Cell Biology, 2009. **185**.
37. Kawaoka, S., et al., *3' end formation of PIWI-interacting RNAs in vitro*. Molecular cell, 2011.
38. Gunawardane, L.S., et al., *A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila*. Science, 2007. **315**(5818): p. 1587-90.

39. Huang, H., et al., *AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification*. J Cell Biol, 2014. **206**(2): p. 217-30.
40. Li, C., et al., *Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies*. Cell, 2009. **137**(3): p. 509-21.
41. Anand, A. and T. Kai, *The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila*. EMBO J, 2012. **31**(4): p. 870-82.
42. Zhang, Z., et al., *Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains*. Mol Cell, 2011. **44**(4): p. 572-84.
43. Xiol, J., et al., *RNA clamping by vasa assembles a piRNA amplifier complex on transposon transcripts*. Cell, 2014. **157**(7): p. 1698-711.
44. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-1299.
45. Klenov, M., et al., *Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline*. Nucleic acids research, 2007. **35**(16): p. 5430-5438.
46. Chambeyron, S., et al., *piRNA-mediated nuclear accumulation of retrotransposon transcripts in the Drosophila female germline*. Proceedings of the National Academy of Sciences, 2008. **105**(39): p. 14964-14969.
47. Shpiz, S., et al., *rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus*. Nucleic acids research, 2009. **37**(1): p. 268-278.
48. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic acids research, 2011. **39**(20): p. 8703-8711.
49. Rozhkov, N., M. Hammell, and G. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes & development, 2013. **27**(4): p. 400-412.
50. Soper, S., et al., *Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis*. Developmental cell, 2008. **15**(2): p. 285-297.
51. Song, S.U., et al., *Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the gypsy retroelement of Drosophila*. Development, 1997. **124**(14): p. 2789-98.



52. al-Mukhtar, K.A. and A.C. Webb, *An ultrastructural study of primordial germ cells, oogonia and early oocytes in Xenopus laevis*. J Embryol Exp Morphol, 1971. **26**(2): p. 195-217.
53. Eddy, E.M. and S. Ito, *Fine structural and radioautographic observations on dense perinuclear cytoplasmic material in tadpole oocytes*. J Cell Biol, 1971. **49**(1): p. 90-108.
54. Mahowald, A.P., *Polar granules of Drosophila. II. Ultrastructural changes during early embryogenesis*. J Exp Zool, 1968. **167**(2): p. 237-61.
55. Hegner, R.W., *The History of the Germ Cells in the Paedogenetic Larva of Miastor*. Science, 1912. **36**(917): p. 124-6.
56. Hegner, R.W., *The Germ Cell Determinants in the Eggs of Chrysomelid Beetles*. Science, 1911. **33**(837): p. 71-2.
57. Hathaway, D.S. and G.G. Selman, *Certain aspects of cell lineage and morphogenesis studied in embryos of Drosophila melanogaster with an ultra-violet micro-beam*. J Embryol Exp Morphol, 1961. **9**: p. 310-25.
58. Okada, M., I.A. Kleinman, and H.A. Schneiderman, *Restoration of fertility in sterilized Drosophila eggs by transplantation of polar cytoplasm*. Dev Biol, 1974. **37**(1): p. 43-54.
59. Aravin, A. and D. Chan, *piRNAs meet mitochondria*. Developmental cell, 2011. **20**(3): p. 287-288.
60. Ipsaro, J.J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-83.
61. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Mol Cell, 2013. **50**(5): p. 749-61.
62. Ponting, C.P., *Tudor domains in proteins that interact with RNA*. Trends Biochem Sci, 1997. **22**(2): p. 51-2.
63. Klattenhoff, C., et al., *The Drosophila HPI homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters*. Cell, 2009. **138**(6): p. 1137-49.
64. Pane, A., et al., *The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline*. EMBO J, 2011. **30**(22): p. 4601-15.

65. Zhang, F., et al., *UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery*. Cell, 2012. **151**(4): p. 871-84.
66. Lim, A.K. and T. Kai, *Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster*. Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6714-9.
67. Patil, V.S. and T. Kai, *Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas*. Curr Biol, 2010. **20**(8): p. 724-30.
68. Anne, J. and B.M. Mechler, *Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuleen*. Development, 2005. **132**(9): p. 2167-77.
69. Kirino, Y., et al., *Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization*. RNA, 2010. **16**(1): p. 70-8.
70. Liu, H., et al., *Structural basis for methylarginine-dependent recognition of Aubergine by Tudor*. Genes Dev, 2010. **24**(17): p. 1876-81.
71. Kirino, Y., et al., *Arginine methylation of vasa protein is conserved across phyla*. J Biol Chem, 2010. **285**(11): p. 8148-54.
72. Liu, L., et al., *PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition*. Development, 2011. **138**(9): p. 1863-73.
73. Hay, B., et al., *Identification of a component of Drosophila polar granules*. Development, 1988. **103**(4): p. 625-40.
74. Anne, J., *Targeting and anchoring Tudor in the pole plasm of the Drosophila oocyte*. PLoS One, 2010. **5**(12): p. e14362.
75. Arkov, A.L., et al., *The role of Tudor domains in germline development and polar granule architecture*. Development, 2006. **133**(20): p. 4053-62.
76. Anne, J., et al., *Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila*. Development, 2007. **134**(1): p. 137-46.
77. Thomson, T. and P. Lasko, *Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning*. Genesis, 2004. **40**(3): p. 164-70.

78. Kugler, J.M. and P. Lasko, *Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis*. Fly (Austin), 2009. **3**(1): p. 15-28.
79. Liang, L., W. Diehl-Jones, and P. Lasko, *Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities*. Development, 1994. **120**(5): p. 1201-11.
80. Megosh, H.B., et al., *The role of PIWI and the miRNA machinery in Drosophila germline determination*. Curr Biol, 2006. **16**(19): p. 1884-94.
81. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline*. Curr Biol, 2001. **11**(13): p. 1017-27.
82. Bozzetti, M.P., et al., *The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2*. Proc Natl Acad Sci U S A, 1995. **92**(13): p. 6067-71.
83. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-760.
84. Vagin, V.V., et al., *A distinct small RNA pathway silences selfish genetic elements in the germline*. Science, 2006. **313**(5785): p. 320-4.
85. Aravin, A.A., et al., *Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line*. Mol Cell Biol, 2004. **24**(15): p. 6742-50.
86. Handler, D., et al., *A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors*. EMBO J, 2011. **30**(19): p. 3977-93.
87. Stapleton, W., S. Das, and B.D. McKee, *A role of the Drosophila homeless gene in repression of Stellate in male meiosis*. Chromosoma, 2001. **110**(3): p. 228-40.
88. Kibanov, M.V., et al., *A novel organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing*. Mol Biol Cell, 2011. **22**(18): p. 3410-9.

89. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(46): p. 18760-5.
90. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-9.
91. Aravin, A.A., G.J. Hannon, and J. Brennecke, *The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race*. Science, 2007. **318**(5851): p. 761-4.
92. Siomi, M.C., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nat Rev Mol Cell Biol, 2011. **12**(4): p. 246-58.
93. Ma, J.B., K. Ye, and D.J. Patel, *Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain*. Nature, 2004. **429**(6989): p. 318-22.
94. Ma, J.B., et al., *Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein*. Nature, 2005. **434**(7033): p. 666-70.
95. Yuan, Y.R., et al., *Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage*. Mol Cell, 2005. **19**(3): p. 405-19.
96. Song, J.J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science, 2004. **305**(5689): p. 1434-7.
97. Lin, H. and A.C. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development, 1997. **124**(12): p. 2463-76.
98. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-60.
99. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-103.
100. Sienski, G., D. Donertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-80.

101. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes Dev, 2013. **27**(4): p. 390-9.
102. Rozhkov, N.V., M. Hammell, and G.J. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes Dev, 2013. **27**(4): p. 400-12.
103. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic Acids Res, 2011. **39**(20): p. 8703-11.
104. Eddy, E.M., *Germ plasm and the differentiation of the germ cell line*. Int Rev Cytol, 1975. **43**: p. 229-80.
105. Mahowald, A.P., *Polar granules of drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis*. J Exp Zool, 1971. **176**(3): p. 345-52.
106. Patil, V.S. and T. Kai, *Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas*. Curr Biol, 2010.
107. Findley, S.D., et al., *Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage*. Development, 2003. **130**(5): p. 859-71.
108. Lim, A.K., L. Tao, and T. Kai, *piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline*. J Cell Biol, 2009. **186**(3): p. 333-42.
109. Nishida, K.M., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. EMBO J, 2009. **28**(24): p. 3820-31.
110. Schupbach, T. and E. Wieschaus, *Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology*. Genetics, 1991. **129**(4): p. 1119-36.
111. Harris, A.N. and P.M. Macdonald, *Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C*. Development, 2001. **128**(14): p. 2823-32.
112. Bossing, T., C.S. Barros, and A.H. Brand, *Rapid tissue-specific expression assay in living embryos*. Genesis, 2002. **34**(1-2): p. 123-6.
113. Djuranovic, S., et al., *Allosteric regulation of Argonaute proteins by miRNAs*. Nat Struct Mol Biol, 2010. **17**(2): p. 144-50.

114. Parker, J.S., S.M. Roe, and D. Barford, *Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex*. Nature, 2005. **434**(7033): p. 663-6.
115. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-7.
116. Snee, M.J. and P.M. Macdonald, *Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components*. J Cell Sci, 2004. **117**(Pt 10): p. 2109-20.
117. Nagao, A., et al., *Gender-Specific Hierarchy in Nuage Localization of PIWI-Interacting RNA Factors in Drosophila*. Front Genet, 2011. **2**: p. 55.
118. Wang, Y., et al., *Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes*. Nature, 2009. **461**(7265): p. 754-61.
119. Aravin, A.A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS Genet, 2009. **5**(12): p. e1000764.
120. Aravin, A.A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Mol Cell, 2008. **31**(6): p. 785-99.
121. Shoji, M., et al., *The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline*. Dev Cell, 2009. **17**(6): p. 775-87.
122. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Mol Cell, 2013. **50**(5): p. 762-77.
123. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nat Cell Biol, 2009. **11**(5): p. 652-8.
124. Xiol, J., et al., *A role for Fkbp6 and the chaperone machinery in piRNA amplification and transposon silencing*. Mol Cell, 2012. **47**(6): p. 970-9.
125. Cora, E., et al., *The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs*. RNA, 2014. **20**(6): p. 773-81.
126. Vermaak, D., S. Henikoff, and H.S. Malik, *Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila*. PLoS Genet, 2005. **1**(1): p. 96-108.

127. Obbard, D.J., et al., *The evolution of RNAi as a defence against viruses and transposable elements*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1513): p. 99-115.
128. Simkin, A., et al., *Recurrent and recent selective sweeps in the piRNA pathway*. Evolution, 2013. **67**(4): p. 1081-90.
129. Comeron, J.M., *A method for estimating the numbers of synonymous and nonsynonymous substitutions per site*. J Mol Evol, 1995. **41**(6): p. 1152-9.
130. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proc Natl Acad Sci U S A, 2011. **108**(46): p. 18760-5.
131. Anne, J., *C-terminal moiety of Tudor contains its in vivo activity in Drosophila*. PLoS One, 2010. **5**(12): p. e14378.
132. Creed, T.M., et al., *Novel role of specific Tudor domains in Tudor-Aubergine protein complex assembly and distribution during Drosophila oogenesis*. Biochem Biophys Res Commun, 2010. **402**(2): p. 384-9.
133. Vagin, V.V., G.J. Hannon, and A.A. Aravin, *Arginine methylation as a molecular signature of the Piwi small RNA pathway*. Cell Cycle, 2009. **8**(24): p. 4003-4.
134. Arkov, A.L. and A. Ramos, *Building RNA-protein granules: insight from the germline*. Trends Cell Biol, 2010. **20**(8): p. 482-90.
135. Siomi, M.C., T. Mannen, and H. Siomi, *How does the royal family of Tudor rule the PIWI-interacting RNA pathway?* Genes Dev, 2010. **24**(7): p. 636-46.
136. Cavey, M., et al., *Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation*. Development, 2005. **132**(3): p. 459-68.
137. Ghabrial, A. and T. Schupbach, *Activation of a meiotic checkpoint regulates translation of Gurken during Drosophila oogenesis*. Nat Cell Biol, 1999. **1**(6): p. 354-7.
138. Klattenhoff, C., et al., *Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response*. Dev Cell, 2007. **12**(1): p. 45-55.

139. Vagin, V.V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes Dev, 2009. **23**(15): p. 1749-62.
140. Bedford, M.T. and S.G. Clarke, *Protein arginine methylation in mammals: who, what, and why*. Mol Cell, 2009. **33**(1): p. 1-13.
141. Krause, C.D., et al., *Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential*. Pharmacol Ther, 2007. **113**(1): p. 50-87.
142. Wang, J., et al., *Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis*. Curr Biol, 2009. **19**(8): p. 640-4.
143. Saxe, J.P., et al., *Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline*. EMBO J, 2013. **32**(13): p. 1869-85.
144. Pandey, R.R., et al., *Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16492-7.
145. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 858-62.
146. Comeron, J.M., *K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals*. Bioinformatics, 1999. **15**(9): p. 763-4.



## MATERIALS AND METHODS

### *Generation of Transgenic Fly Lines*

Transgenic protein constructs for injection were generated using Gateway cloning system (Life Technologies). Genes were obtained by PCR from *Drosophila melanogaster* template cDNA taken from ovaries or testes of adult flies. Point mutations were engineered by overlap PCR and inserted in the pENTR-D-TOPO vector. Transgenes were cloned into the pUASP-Gateway-phiC31 fly injection vector derived from pCasPeR5-phiC31 vector containing GFP, mKate2, or Strep-FLAG tags. All transgenes created for this manuscript contain N-terminal tags, with the exception of SpnE and Zuc transgenes, which were C-terminally tagged. The expression of each transgene was controlled using the yeast upstream activation sequence promoter (UASp) coupled with a maternal  $\alpha$ -Tubulin67c-Gal4-VP16 (MaG4) driver, which showed strong expression in nurse cells starting at stages 2-4 of oogenesis onwards to the oocyte (Figure 1 – figure supplement 1A). Transgenes were generated in flies by PhiC31-mediated transformation (BestGene) using PhiC31 landing pads on either chromosome two (BDSC #9736) or chromosome three (BDSC #9750).

The GFP-Aub BAC line was generated by cloning of the *aub* genomic locus from the BAC clone BACN04M10 into the pCasPeR4 vector using restriction sites XhoI and SpeI. Bacterial recombineering (Gene Bridges Counter Selection kit) was used to insert an in-frame GFP tag in the start site of Aub.

### *Cell Culture, Co-immunoprecipitation, & Western Blots:*

Schneider S2 cells were cultured in complete Schneider medium (10% heat inactivated FBS; 100 units of penicillin [Life technologies]; 100µg streptomycin [Life technologies]). Plasmids were generated using Gateway cloning technologies (Life technologies) using the Drosophila Gateway Vector Collection (DGVC) destination vectors, pAGW for GFP and pAFW for 3xFLAG tags, expressed by the Actin5C promoter. Cells were transfected using TransIT-LT1 transfection reagent (Mirus biosciences) according to protocol but

with 3.0µg of total plasmid, with 1.5µg each plasmid DNA containing 3xFLAG or GFP-tagged genes.

For co-immunoprecipitation of proteins expressed in S2 cells, one well of a six-well tissue culture plate of transfected S2 cells were carefully lysed in 150µL S2 Lysis buffer (20mM Tris at pH7.4, 150 KCl, 0.1% Tween-20, 0.1% Igepal, EDTA-free Complete Protease Inhibitor Cocktail [Roche]) and incubated for 10 minutes on ice. Supernatant was cleared by centrifugation at 4,000 x g for 10 minutes at 4°C. Input sample was collected from the supernatant. Anti-FLAG M2 beads (Sigma Aldrich) were blocked in 5mg/ml BSA for 10min at 4°C, followed by washing in S2 lysis buffer. Beads were added to the supernatant and rotated at 4°C for 90min. Beads were washed three times in PBS + 0.05% Tween-20 and eluted by boiling in reducing SDS loading buffer.

#### *piRNA sequencing & Analysis*

200 ovaries from flies fed yeast were lysed in lysis buffer (20mM HEPES at pH 7.0, 150mM KCl, 2.5mM MgCl<sub>2</sub>, 0.5% Triton X-100, 0.5% Igepal, 100 U/mL RNasin [Promega], EDTA-free Complete Protease Inhibitor Cocktail [Roche]) and supernatant clarified by centrifugation at 15,000x g. Supernatant was incubated with anti-eGFP polyclonal antibody (Covance) conjugated to Protein-G Dynabeads at 4°C and washed three times in NT2 buffer (see: *Co-immunoprecipitation & Western Blots*). Beads were spiked with 5 pmol of synthesized 42-nt (42M) RNA oligomer to assess purification efficiency (see: *Semi-Quantitative piRNA Binding Affinity Analysis*), proteinase K-digested, and phenol-extracted. Isolated RNA was CIP-treated, radiolabeled using PNK and gamma-P32-labeled ATP, and run on a 15% urea-PAAG gel. Small RNAs were cloned according to published methods [99, 145].

#### *Semi-Quantitative piRNA Binding Affinity Analysis*

Normalized Signal Intensities of piRNA levels after IP of GFP-tagged Aub and Ago3 transgenes was achieved by a normalized semi-quantitative approach. The gel band intensity corresponding to 29-19nt piRNA size range from autoradiographs ( $I_{piRNA}$ ) were normalized to western blot band intensities corresponding to GFP-tagged transgenes after IP ( $I_{protein}$ ). To account for potential losses in piRNA yield due to RNA purification, IP

were spiked with 5pmol of synthesized 42-nt (42M) RNA oligomer prior to proteinase K digestion and phenol extracton. Autoradiographed gel band intensities from the 42M oligomer ( $I_{42M}$ ) were subtracted from western blot band intensities. Background signal intensities from adjacent lane regions of equal area were subtracted from each band intensity measurement. These values were then used to obtain the piRNA normalized signal intensity using the following calculation:

$$piRNA \text{ normalized signal intensity} = \frac{[I_{piRNA} - I_{BGpiRNA}]}{[I_{protein} - I_{BGprotein}] - [I_{42M} - I_{BG42M}]}$$

Band intensities were obtained using FIJI software and values graphed using Microsoft Excel.

#### *Divergence Analysis*

cDNA sequences for *D. melanogaster* and *D. simulans* were aligned in PRANK using translated codons as reference and aligned again using ClustalW2 after removing indels, stop codons and unalignable sequences (PRANK: <http://www.ebi.ac.uk/goldman-srv/prank/prank>)(CLUSTALW2: <http://www.ebi.ac.uk/Tools/msa/clustalw2>). K-estimator software was used to determine dN and dS values arbitrarily using a 100bp window and 35bp step size [126, 129, 146] (<http://www.biology.uiowa.edu/labs/comeron/software>). Values obtained after divergence analysis were graphed using Microsoft Excel.

#### *Microscopy*

Ovaries were fixed in 4% PFA in PBS for 20 minutes, permeabilized in 1% Triton-X100 in PBS, and DAPI stained (Sigma-Aldrich). Actin was stained in specimen as indicated in the figure legends with Rhodamine Phalloidin (Life Technologies) according to the manufacturer's protocol. Ovaries were washed in PBS and mounted in Vectashield medium (Vector Labs). S2 cells were allowed to settle on coverslips treated with Poly-L-Lysine (Sigma-Aldrich). After gentle washing, cells were fixed in 0.5% PFA in PBS for 20 minutes followed by staining with DAPI (Sigma-Aldrich), washed, and mounted in

Vectashield medium (Vector Labs). Images were captured using an AxioImager microscope; an Apotome structured illumination system was used for optical sections (Carl Zeiss).

### *Fluorescence Recovery After Photobleaching (FRAP)*

For each construct, at least 20 independent FRAP experiments using ovaries expressing a single GFP-tagged transgene under the MaG4 driver were performed. FRAP experiments were captured on a Zeiss LSM710 confocal microscope (Carl Zeiss AIM) equipped with a 25x/0.8 NA Imm Corr multi-immersion objective lens and operating Zeiss Zen Black software. Image acquisition for all experiments utilized an identical 488nm AOTF laser power setting of 7% to ensure laser power was not influencing measurements. PMT gain settings were variably set to accommodate the expression level of each GFP-tagged protein. Images were acquired at 256x256 pixel resolution at 0.07µm pixel size and scan speed of 614.4ms per frame with 1.0 µs pixel dwell time. A single bleach region was defined for each experiment, consisting of a region of 7x7 pixels equal to 0.49 µm x 0.49 µm and was bleached by a single iteration of 100% laser power from 488, 561 and 633nm wavelengths. Five initial pre-bleach images were captured prior to bleaching and 115 subsequent post-bleach images were acquired every 614.4ms to assess fluorescence recovery in the bleach zone. FIJI Is Just ImageJ (FIJI; <http://fiji.sc/>) software was used to analyze FRAP experiments. To account for background and photo-bleaching effects during acquisition, the mean intensity values from the bleach zone (*BL*), the background zone (*BG*), and the reference signal zone (*REF*) were used to calculate the corrected BL (*BL<sub>corr</sub>*) for each acquisition frame using the equation:

$$BL_{corr} = \frac{[BL - BG]}{[REF - BG]}$$

*BL<sub>corr</sub>* values were normalized to the mean of five pre-bleach values, which were used to estimate 100% fluorescence intensity. Using the curve fitter module of FIJI, the normalized post-bleach data was fit to an exponential recovery model:

$$y = a(1 - e^{-bx}) + c$$

The mobile fraction was obtained by the sum of coefficients  $a$  and  $c$ , which describes the maximal extent of recovery for each experiment. The mean mobile fraction for each line was calculated from at least 20 replicate FRAP experiments taken from at least four animals from the same stable line.

## FIGURES AND FIGURE LEGENDS

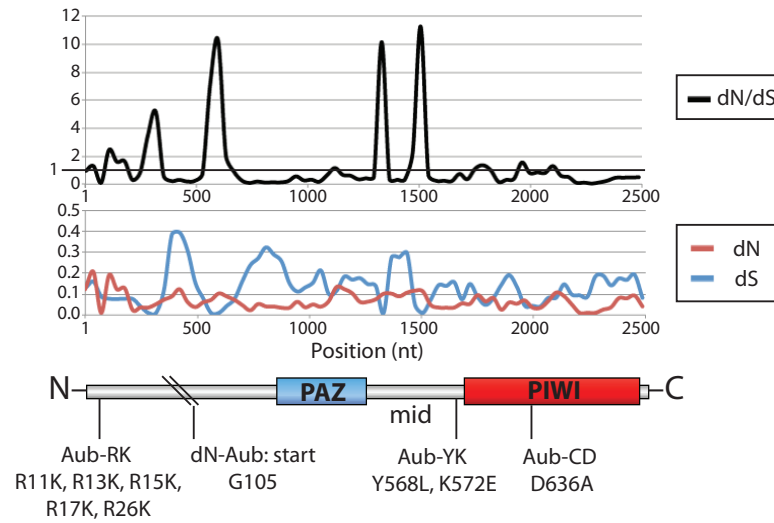
### ***Figure 1: Identity and Divergence analysis of Drosophila Piwi-clade Argonaute proteins.***

(A) ClustalW2 alignment shows the N-terminal region of *D. melanogaster* Aub, Ago3, and Piwi protein sequences to be the least similar region of the proteins. The N-terminal region was divided in two parts (N-term 1 and N-term 2) according to the site of cleavage of dN-Aub and N-Ago3-Aub transgenic proteins. Other domains were defined using PFAM protein family database search criteria. The percentage of amino acids with identity or with strong or weak conservation was obtained by counting ClustalW2 scores of conservation and dividing by the average number of amino acids in each domain of Aub, Ago3, and Piwi sequences. (B) The ratio of synonymous (dS) versus non-synonymous (dN) mutations of *D. melanogaster* compared to *D. simulans* Aub sequence shows moderate elevation in the N-terminal region. (C) The ratio of synonymous (dS) versus non-synonymous (dN) mutations of *D. melanogaster* compared to *D. simulans* Ago3 sequence shows slight elevation in the N-terminal region, although we could not obtain a full length sequence for *D. simulans* Ago3. In general, ratios exceeding 1 are indicative of positive selection, while those less than 1 are indicative of negative selection. Ratios equal to 1 are subject to neutral drift. Values were obtained using 100nt sliding window, with a 35nt step size in K-estimator software. The approximate location of protein domains and point mutations used to create transgenic proteins in this manuscript are indicated below the divergence analysis, along with the name of the transgene described in the main text.

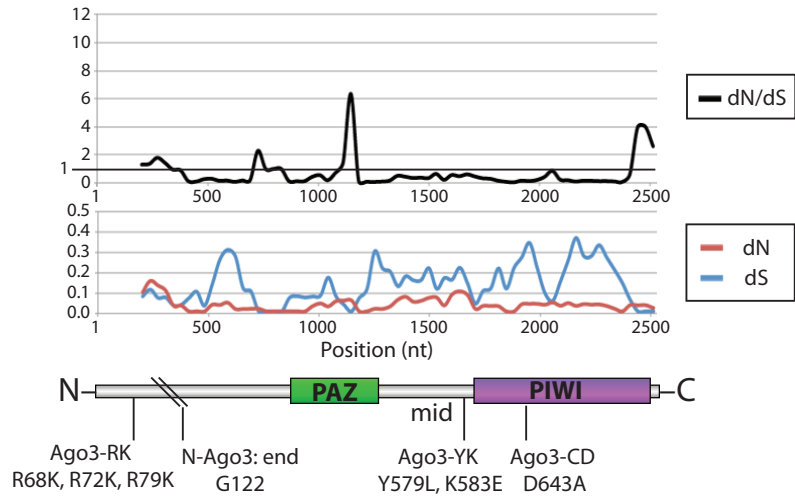
A

Clustal W2 Alignment	N-term 1	N-term 2	PAZ	mid	PIWI
<b>AA Length</b> (Aub; Ago3; Piwi)	105; 122; 91	175; 168; 171	130; 132; 130	142; 143; 144	314; 302; 307
<b>Identical Amino Acids</b> (*)	5% (5)	18% (31)	27% (35)	14% (20)	28% (85)
<b>AA w/ Strong Conservation</b> (.)	10% (11)	27% (46)	27% (35)	23% (33)	25% (77)
<b>AA w/ Weak Conservation</b> (.)	14% (15)	6% (10)	9% (12)	19% (27)	13% (39)

B

Comparison of *D. melanogaster* and *D. simulans* Aubergine (Aub)

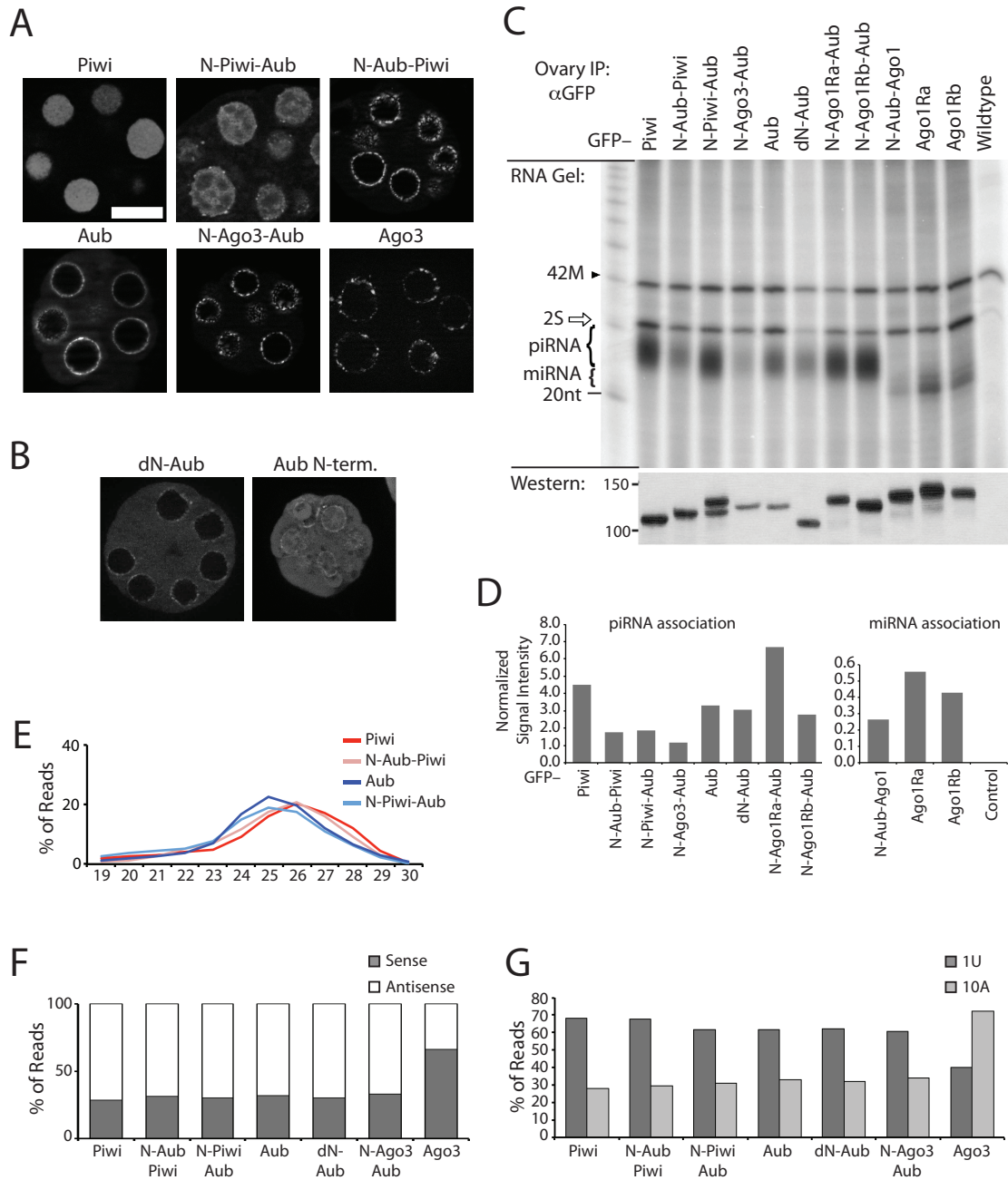
C

Comparison of *D. melanogaster* and *D. simulans* Argonaute-3 (Ago3)

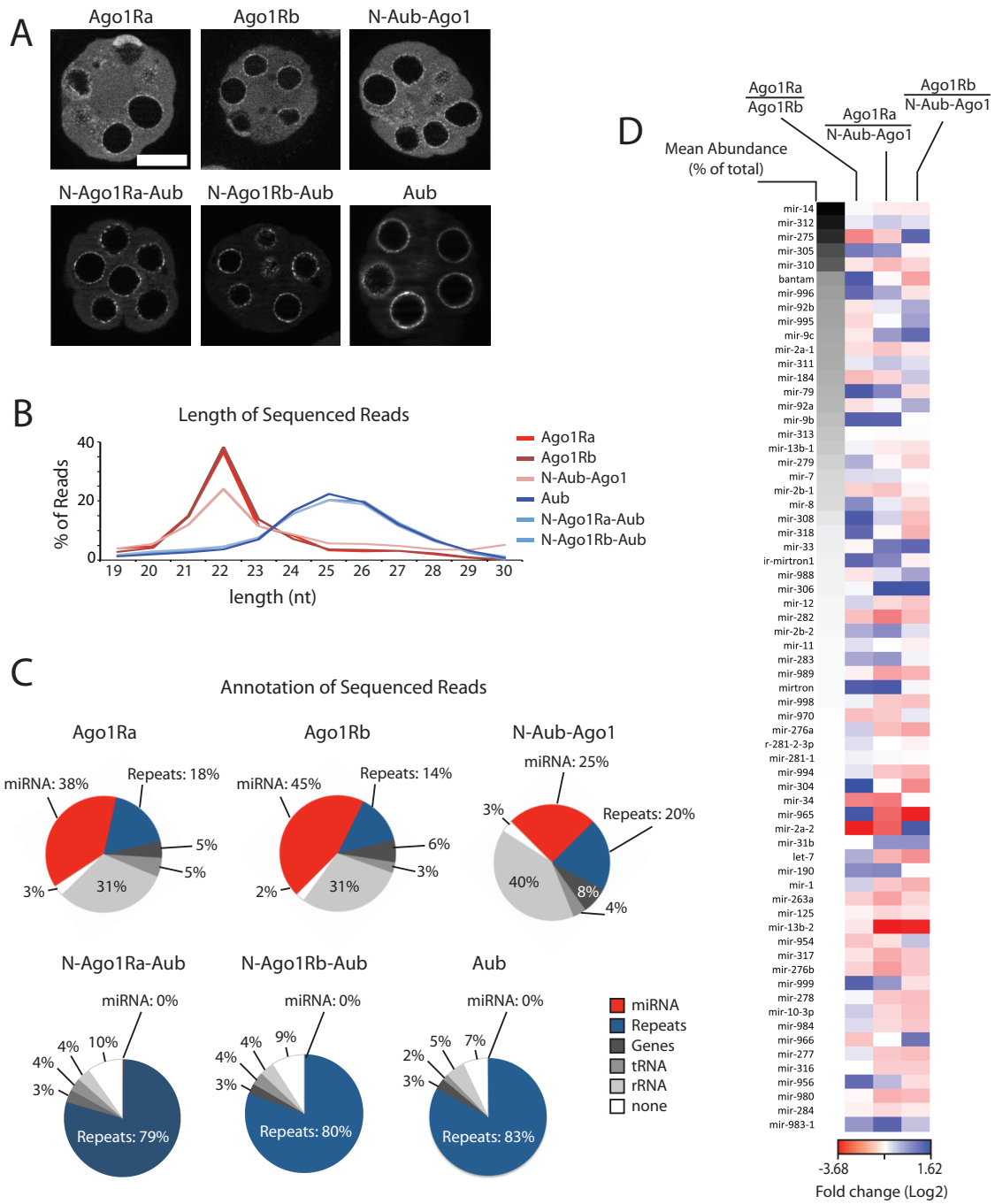
***Figure 2: N-terminal domain swaps reveal localization and piRNA loading of Piwi, Aub, and Ago3 are dependent on PAZ-mid-PIWI domains.***

(A) Localization of GFP-tagged chimeric domain swapped Piwi proteins in ovary nurse cells reveal that N-terminal region of Piwi functions as a nuclear localization signal. Other proteins localize similarly to Aub and Ago3 to perinuclear nuage. Scale bar is 10µm. (B) Localization of GFP-tagged N-terminally truncated Aub shows localization to nuage and diffused throughout the cytosol, while Aub N-terminal region is dispersed throughout the cytoplasm and nucleus of nurse cells, but does not enrich in nuage. (C) RNA gel of associated small RNA co-purified with GFP-tagged chimeric proteins. 42M indicates a synthetic 42nt RNA oligomer used for normalization in quantifying RNA relative to immunoprecipitated protein levels. A band at 30nt is 2S ribosomal RNA that co-purifies non-specifically with all immunoprecipitations. The size ranges for piRNA (24-29 nt) and miRNA (22nt) are indicated. (D) Semi-quantitative analysis of associated piRNA and miRNA from N-domain swapped constructs. (E) Length distribution of piRNA sequences co-purified from N-domain swapped constructs. (F) Proportion of sense to antisense piRNA sequences co-purified with N-domain swapped constructs shows that all proteins predominantly associated with anti-sense sequences, with the exception of Ago3. (G) Nucleotide bias of piRNA sequences co-purified with N-domain swapped constructs. All proteins have a bias for 1U and not 10A, with the exception of Ago3.





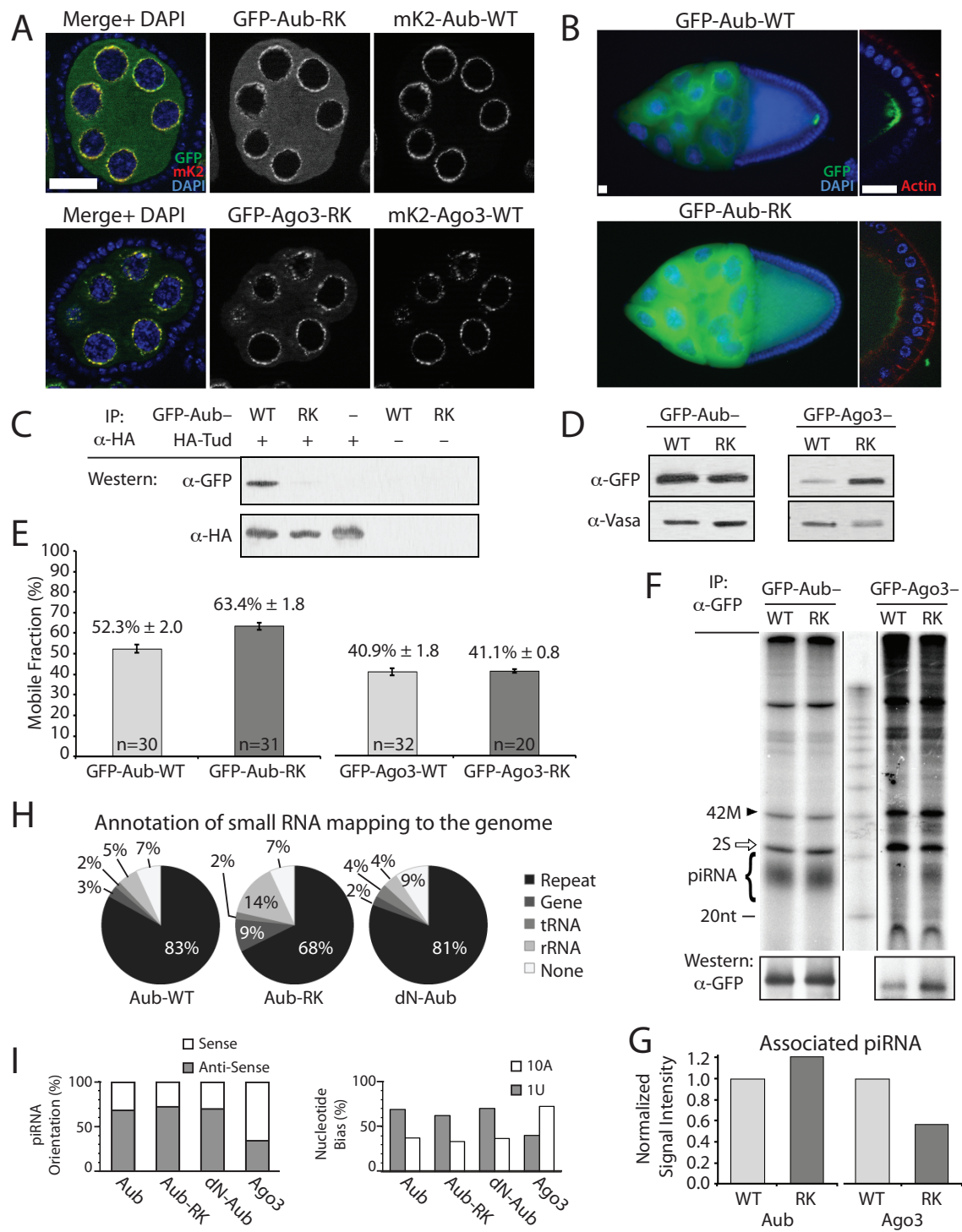
**Figure 3: Reciprocal N-terminal domain swaps between Ago1 and Aub confirm the specificity of PAZ-mid-PIWI domains in piRNA and miRNA loading and localization.** (A) Localization of Ago1 and Aub N-terminal domain swaps. (B) Length distribution of small RNA associated with Ago1 and Aub N-terminal domain swaps. (C) Annotation of sequenced reads from N-terminal domain swaps shows that the identity of the PAZ-mid-PIWI domains— and not N-terminal regions— contribute to loading of small RNAs. (D) Heatmap comparison of miRNA levels associated with two Ago1 isoforms (Ago1Ra and Ago1Rb) and the chimeric N-Aub-Ago1 protein. The heatmap is ranked from highest to lowest by mean abundance of miRNA co-purified from ovaries expressing one of the three proteins.



**Figure 4. Arginine methylation has a minor role in Aub or Ago3 nuage localization, mobility, and the loading with piRNAs.**

(A) The localization patterns of methylation-deficient Aub (GFP-Aub-RK; green; upper panel) and Ago3 (GFP-Ago3-RK; green; lower panel) proteins in nurse cells co-expressed alongside their respective wild-type proteins tagged with mK2 for comparison (red). DNA is stained with DAPI (blue). Scale bar: 20 $\mu$ m. (B) Wild-type Aub accumulates at the pole plasm of the oocyte in late stage egg chambers, while accumulation of Aub-RK is greatly reduced. Actin is stained with rhodamine-phalloidin (red) and DNA is stained with DAPI (blue). Scale bars: 20 $\mu$ m. (C) Wild-type, but not arginine-methylation deficient Aub mutant (Aub-RK) co-immunoprecipitates with Tud. Each tagged pair of proteins was co-expressed in *Drosophila* ovaries. (D) Methylation-deficient Ago3-RK is more abundant relative to its wild-type counterpart. Protein levels in lysate from ovaries expressing either GFP-Aub-RK or GFP-Ago3-RK are compared to their wild-type counterparts using the germ cell specific protein Vas as a loading control. All transgenes are expressed using the same driver. (E) A deficiency in arginine methylation increases the fraction of mobile Aub protein. The mobile fraction of nuage protein was determined for Aub-RK and Ago3-RK mutants in replicate FRAP experiments ( $n=31$  and  $n=20$ , respectively). The fraction of mobile Ago3-RK remains similar to that of wild-type Ago3 in nuage. Bars show standard error values for each set of experiments. (F) Arginine methylation of Aub and Ago3 does not affect the levels of piRNAs co-purifying with each protein. piRNAs were isolated from immunoprecipitated wild-type and methylation-deficient Aub and Ago3 mutants and 5' labeled with P32-containing phosphate. piRNAs are observed between 20 and 30 nt markers. The strong band at 30 nt is abundant 2S ribosomal RNA (white arrow). An unlabeled 42 nt spike RNA (42M) was spiked into each IP sample to control for RNA loss during isolation and labeling efficiency (arrowhead). The levels of each protein after immune-purification are shown by western blot (lower panels). An equal fraction of immuno-purified material was loaded in each lane. (G) Semi-quantitative piRNA-binding analysis comparing the signal intensity of radiolabeled piRNA in Figure 2F shows that GFP-Aub-RK co-purifies with more piRNA than GFP-Aub-WT, while GFP-Ago3-RK has less piRNA than the wildtype transgenic protein. The signal of radiolabeled piRNA was normalized to the amount of immunoprecipitated GFP-tagged transgenic Aub or Ago3 protein and to the

amount of 42nt marker (42M), which was used to normalize against losses during RNA purification. (H) The annotation of piRNAs associated with wild-type Aub, Aub-RK, and dN-Aub show little difference in piRNA content. In this analysis, only those sequences greater than 22 nt but less than 29nt in length are shown. (I) The orientation and nucleotide biases of piRNAs associated with Aub-RK, and dN-Aub are most similar to Aub-WT piRNAs. Only those sequences greater than 22 nt but less than 29nt in length mapping to annotated repeats are analyzed.



***Figure 5 Aub localizes to nuage even in the absence of PRMT5 methylosome components, Csul and Vls.***

The nuage localization of GFP-Aub-WT appears mostly unperturbed in PRMT5 methylosome complex mutant backgrounds for Capsul  n (Csul) (A) or Valois (Vls) (B). An increase in cytoplasmic GFP-Aub-WT is apparent in these mutant backgrounds that is similar to the localization of the GFP-Aub-RK transgenic protein. DNA is stained with DAPI (blue). Scale bar: 20  m.

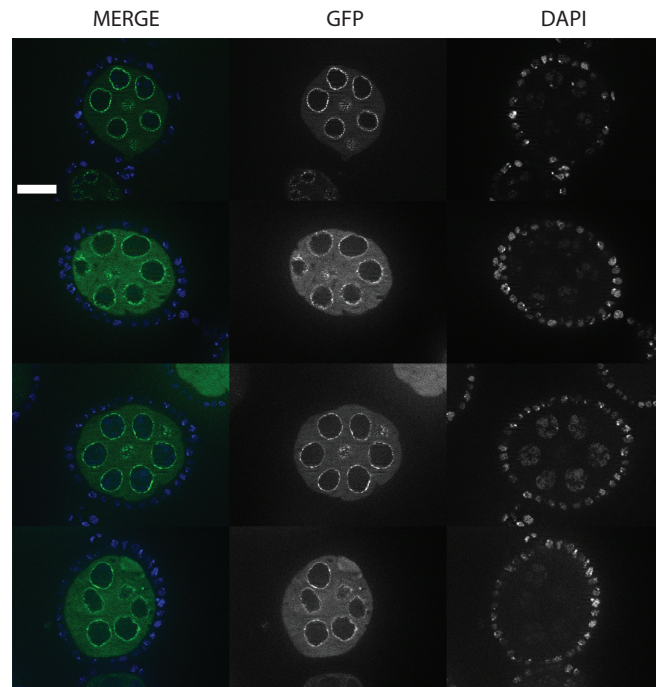
A

$\frac{\text{Csul [RM50]}}{\text{CyO}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-WT}}{\text{TM6}}$

$\frac{\text{Csul [RM50]}}{\text{Csul [5.1]}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-WT}}{\text{TM6}}$

$\frac{\text{Csul [RM50]}}{\text{CyO}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-RK}}{\text{TM6}}$

$\frac{\text{Csul [RM50]}}{\text{Csul [5.1]}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-RK}}{\text{TM6}}$

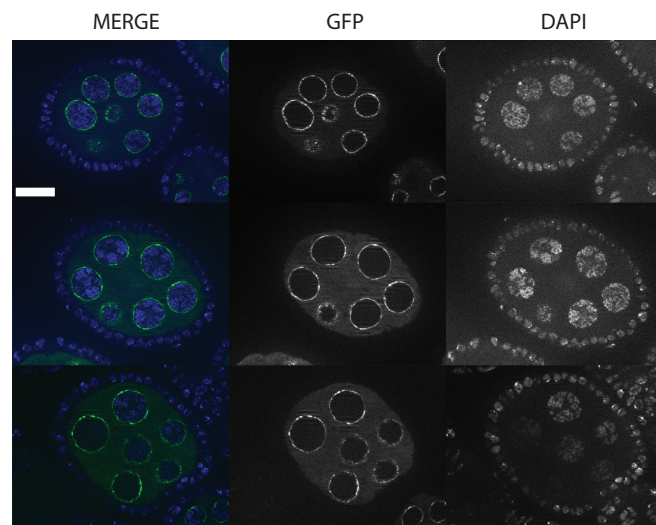


B

$\frac{\text{Vls [1]}}{\text{CyO}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-WT}}{\text{TM6}}$

$\frac{\text{Vls [1]}}{\text{Vls [2]}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-WT}}{\text{TM6}}$

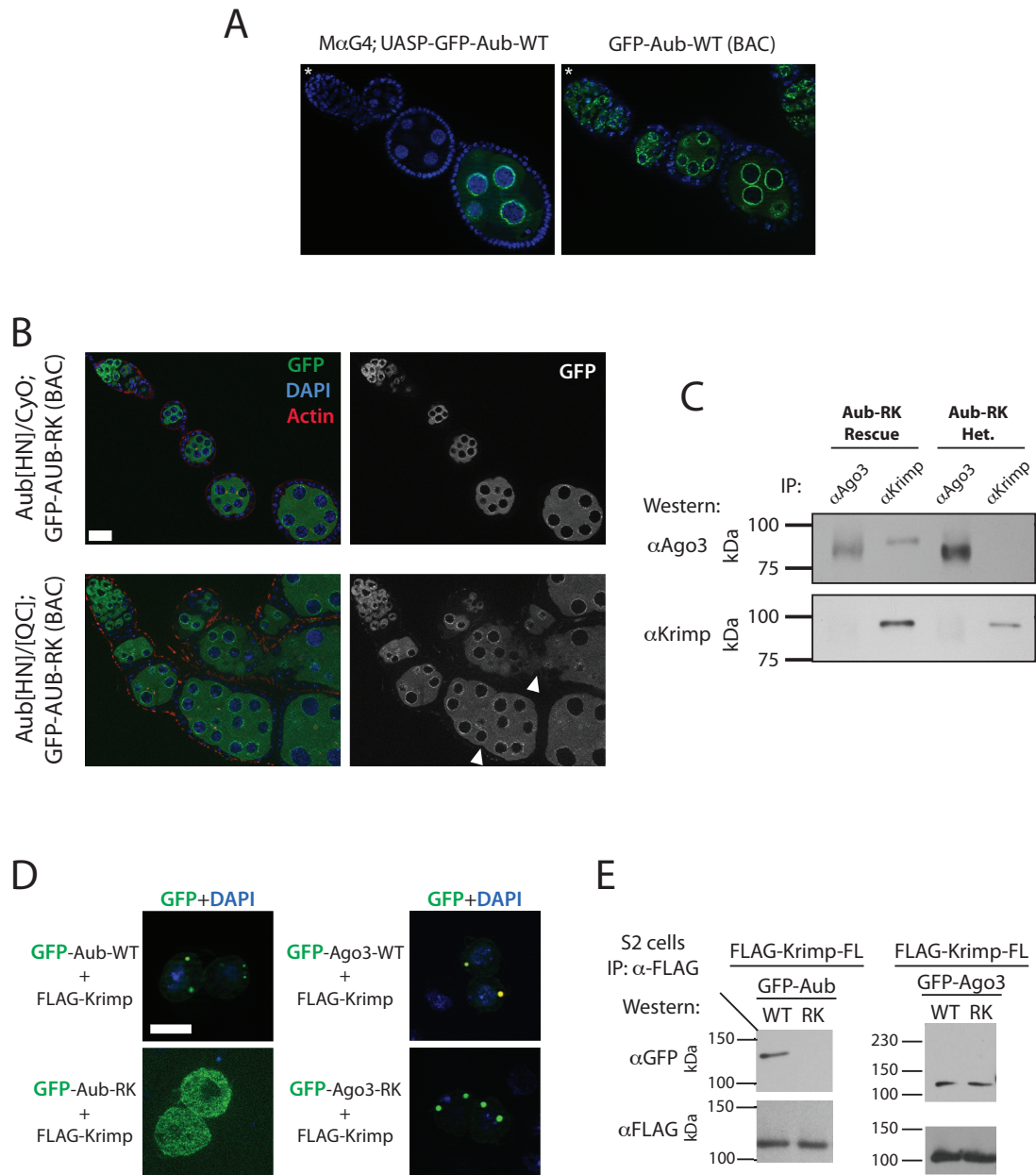
$\frac{\text{Vls [1]}}{\text{CyO}} ; \frac{\text{M}\alpha\text{G, UASP-GFP-Aub-RK}}{\text{TM6}}$





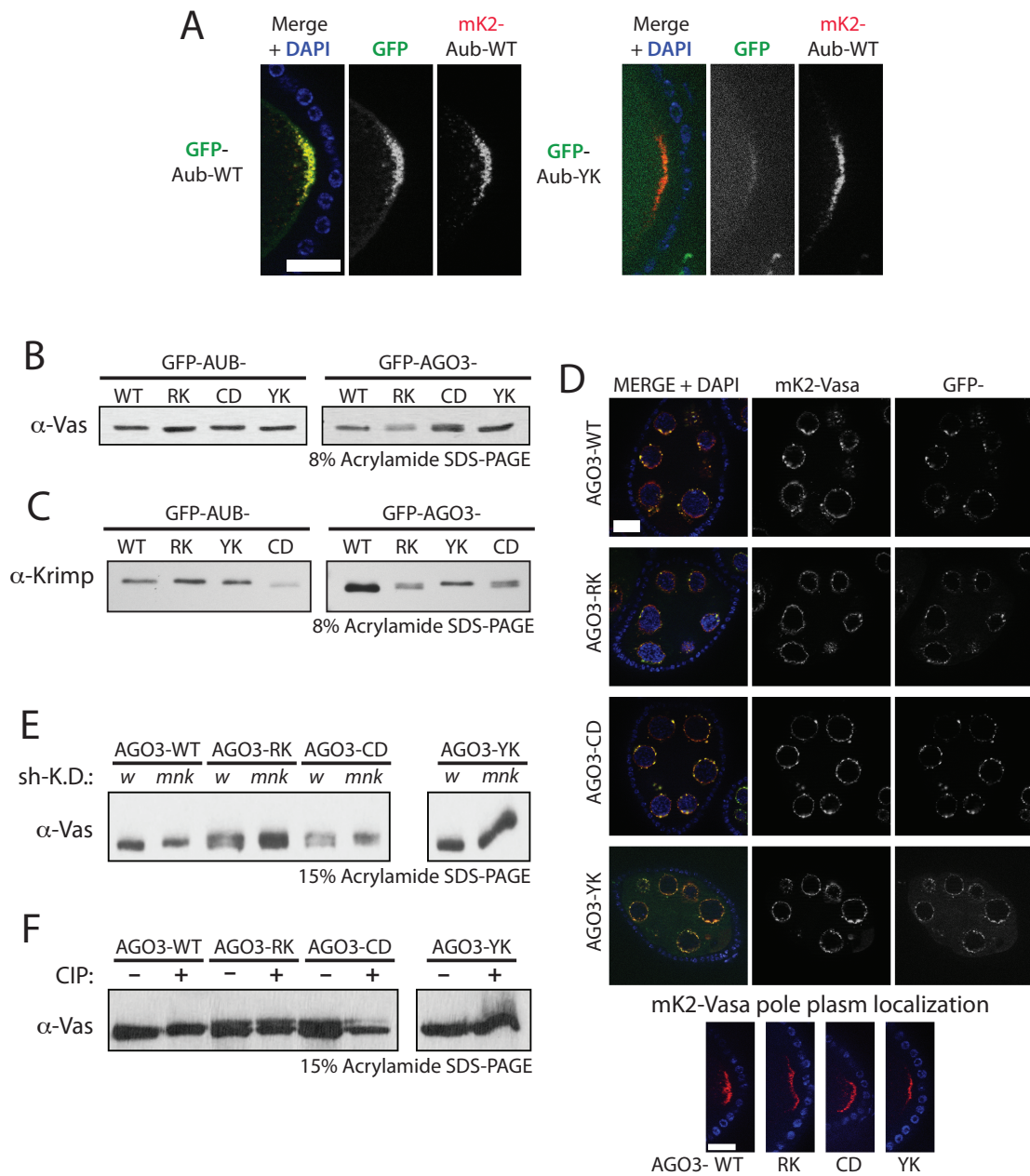
***Figure 6: Arginine methylation of Aub is required for fertility and interactions with Krimp.***

(A) Expression of MaG4 driven UASp-GFP-Aub-WT (left panel; green) begins at stage 4-6 of oogenesis, while a GFP-Aub-WT BAC line (right panel; green) expressing transgenic protein under the native Aub promoter begins at stage one of the germarium and continues onward to the oocyte. The star at the top left of each image marks the anterior tip of the germarium for each line. DNA is stained with DAPI (blue). (B) Representative images showing morphological defects of Aub heterozygous (upper panel) and mutant (lower panel) ovaries expressing GFP-tagged methylation-deficient Aub protein expressed under its native regulatory elements. (C) Co-immunoprecipitation of Ago3 or Krimp in Aub-RK rescue ovaries or heterozygous litter mates demonstrates the stabilization of Ago3-Krimp complex in the absence of Aub methylation. (D) S2 cells transfected with GFP-tagged Aub protein or methylation deficient Aub-RK in the presence of FLAG-tagged Krimp reveal that methylation is required for granule formation. Conversely, Ago3 associates with Krimp granules independently of arginine methylation. (E) Co-immunoprecipitation experiments in S2 cells demonstrate that Aub methylation is required for it to interact with Krimp. Ago3 interacts with Krimp independently of arginine methylation.



**Figure 7: piRNA binding-deficient mutant Aub protein, Aub-YK, does not localize to the pole plasm.**

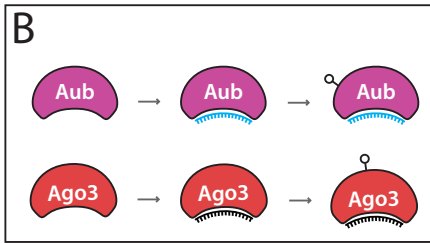
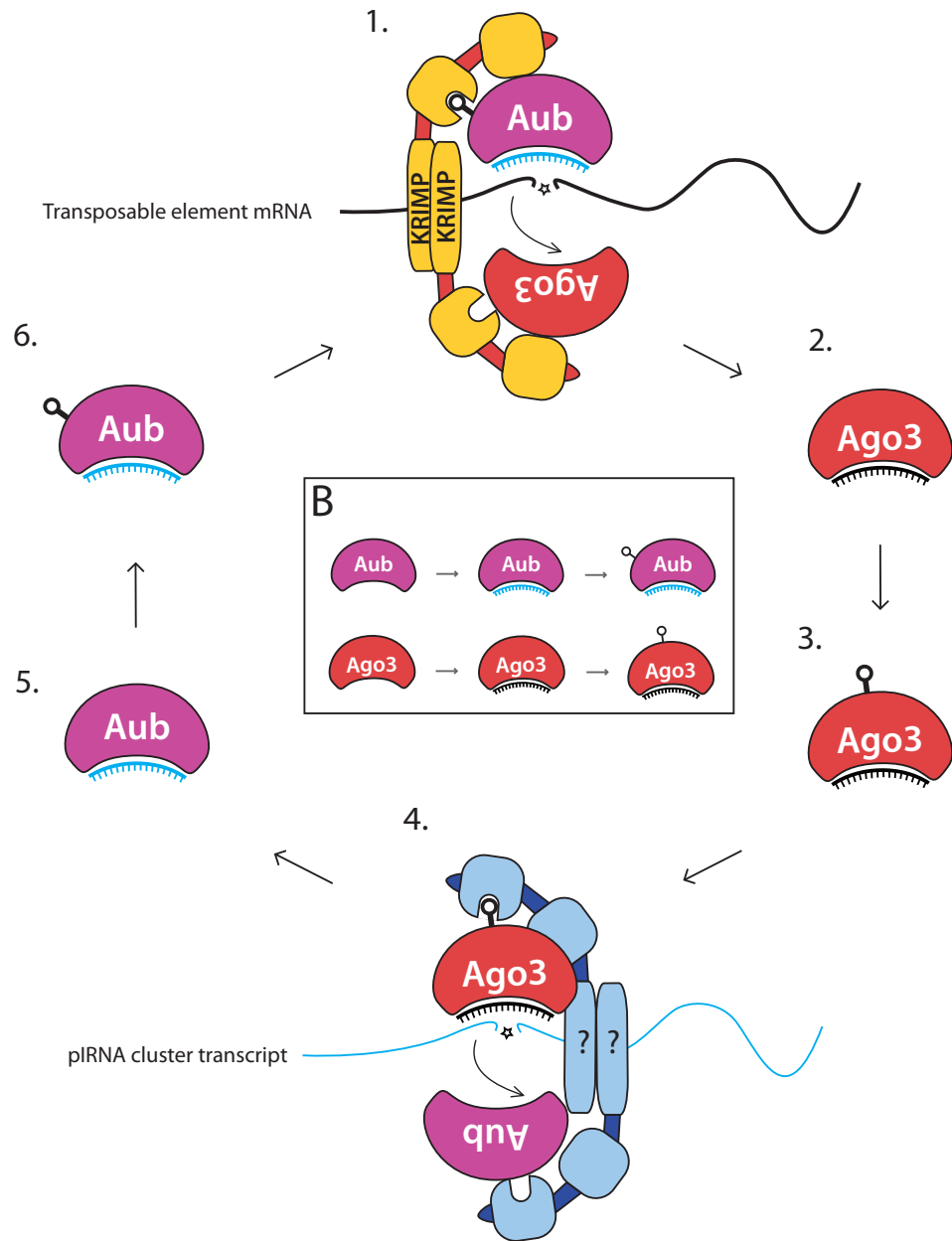
(A) The pole plasm localization of GFP-Aub-YK (green, lower panel) is mostly absent compared to mKate2-Aub-WT (red) or GFP-Aub-WT (green, upper panel). DNA is stained with DAPI (blue). Scale Bar: 20µm. (B) Western blot of Vasa protein from lysate of transgenic fly lines expressing GFP-tagged Aub and Ago3 proteins reveals the presence of two isoforms of different mass in ovaries expressing methylation deficient (RK), or catalytically inactivated (CD) proteins of Ago3. Transgenic fly lines expressing GFP-tagged Aub protein that are wildtype (WT), methylation deficient (RK), catalytically inactivated (CD), or piRNA binding deficient (YK) do not show accumulation of modified Vasa isoform, nor do Ago3-WT or piRNA binding deficient Ago3-YK. (C) Western blot of Krimp protein from lysate of transgenic fly lines expressing GFP-tagged Aub and Ago3 proteins reveals the presence of two isoforms of different mass in ovaries expressing methylation deficient (RK), or catalytically inactivated (CD) proteins of Ago3. (D) Localization of mKate2-Vas (red) is similar to wildtype in ovaries expressing wildtype (WT), methylation deficient (RK), catalytically inactivated (CD), or piRNA binding deficient (YK) versions of GFP-tagged Ago3 protein (Green), indicating that the modified Vas isoform is not disrupted in nuage localization (top panels) or pole plasm localization (Lower panels). Ovaries are stained for DAPI. Blue Scale bar: 20µm. (E) Knockdown of *Drosophila* Chk2 homolog, *mnk*, does not ‘rescue’ the accumulation of modified vasa isoform. (F) Treatment of lysate with calf intestinal phosphatase (CIP) appears to reduce the levels of modified Vas isoform in lysate from ovaries expressing endonucleolytic impaired Ago3 protein, but not methylation deficient Ago3-RK.



***Figure 8: A model for arginine methylation regulating Aub and Ago3 in the piRNA pathway.***

(A) The model for the Aub-Krimp-Ago3 triple complex presented in Chapter 3 is shown in step 1. In this chapter, we show that Aub must be methylated to interact with Krimp, while Ago3 devoid of piRNA interacts with Krimp independently of arginine methylation. In this complex, Aub targets and cleaves complementary transcripts that are then loaded into Ago3 (1). Loading causes dissociation of the Ago3-Krimp complex (2). Structural changes in Ago3 protein allow N-terminal arginine residues to be methylated by the PRMT5 methylosome complex (3). Once methylated, Ago3 interacts with specific Tudor domains in proteins, including PAPI (4). In this complex, Ago3 targets and cleaves complementary transcripts, which are then loaded into Aub protein (5). Aub loaded with piRNA can then be methylated, enabling it to bind Krimp or other Tudor domain proteins, regenerating the cycle (6). (B) The cycle of Aub and Ago3 methylation is governed by a simple principle, where Aub and Ago3 are only methylated after they are loaded with piRNA.

A



## CHAPTER 5

### **Aub endonucleolytic function stabilizes an intermediate complex containing Piwi.**

This chapter will be published as an individual manuscript.

## INTRODUCTION

In sexually reproducing animals, the potential for offspring is provided by totipotent germ cells. Germ cells must be able to differentiate into all cells necessary to make a new individual of either female or male genders. The genomes of germ cells are therefore in an ‘open’ state where epigenetic marks that specify differentiation of somatic cell types are mostly absent. In this permissive genomic environment, deleterious genetic elements such as transposable elements (TE) that replicate autonomously pose a serious threat to the integrity of the germ cell genome [91].

The piRNA pathway functions to extinguish the replicative potential of TE sequences by destroying targeted transcripts in the cytoplasm and by guiding the deposition of repressive chromatin marks at genomic loci expressing TE transcripts [38, 99, 101-103, 147, 148]. In *Drosophila* ovaries, three Piwi proteins, Piwi, Aubergine (Aub), and Argonaute-3 (Ago3), function non-redundantly to extinguish the expression of TE sequences [38, 40, 97, 99, 111]. Piwi proteins together with their respective piRNA guides form piRNA-Induced Silencing Complexes (piRISC), which are the main effectors of TE repression [38, 39, 99, 101, 103]. Once loaded with piRNA, Piwi-piRISC enters the nucleus to coordinate the co-transcriptional deposition of the repressive chromatin mark, H3K9me3, at targeted loci [100, 101, 103]. In the cytoplasm, Aub- and Ago3-piRISC complexes scan and endonucleolytically cleave TE transcripts that have complementarity to their piRNA guides in the piRNA ping-pong pathway [38, 40, 91, 111].

While the nuclear and cytoplasmic branches of the piRNA pathway are both required for effective TE repression, the relationship between both pathways is not well understood. While Aub and Piwi localize to different compartments of germ cells, they are both loaded with primary piRNA that is predominantly anti-sense to TE sequences and originate from genomically encoded piRNA clusters. Conversely, piRNA loaded into Ago3 is generated after Aub targets and cleaves sense-strand TE mRNA [38, 99]. Secondary piRNA are believed to function in selecting sequences in piRNA cluster transcripts that have better specificity for actively expressed TE sequences [38, 91, 99, 149]. Somehow through this mechanism, ping-pong between Aub and Ago3 should also convey the sequences of piRNA that work best at targeting active TE sequences. Here,



we show for the first time that Aub and Piwi interact in the cytoplasm of germ cells in a manner dependent on Aub endonucleolytic activity. This interaction involves two piRNA-related factors, Spindle-E (SpnE), which is a Tudor domain-containing putative DExD-box RNA helicase, and Maelstrom (Mael), which is a High Mobility Group (HMG) box protein with diverse functions in *Drosophila* oogenesis [20, 107, 150]. Similar to Piwi, both SpnE and Mael localize to the nucleus, but are also abundant in perinuclear nuage alongside Aub and Ago3 proteins. Overall, we believe to have found the first evidence of cross-talk between the cytoplasmic and nuclear branches of the piRNA pathway.

## RESULTS

### *Endonuclease activity is required for the mobility of Aub and Ago3 in nuage*

Aub and Ago3 are able to cleave RNA targets complementary to their guide piRNAs using a conserved RNase H-like fold in their PIWI domains [38, 39, 95, 96]. We addressed the role of the endonuclease activity of Aub and Ago3 proteins in their localization and function by generating a point mutant of the conserved DDH residues of the PIWI domain to catalytically inactivate Aub and Ago3 (Aub-CD and Ago3-CD).

The protein stabilities of catalytically impaired Aub and Ago3 were similar their wild-type counterparts (Figure 1E). Immunoprecipitation of wild-type and catalytically inactive Aub and Ago3 protein revealed that Aub-CD associated with slightly less quantities of piRNA as wild-type protein (Figure 1C). Similar to Aub, an even smaller fraction of Ago3-CD was loaded with piRNAs compared to the wild-type Ago3 (Figure 1D). The localization of catalytically inactive Ago3 was indistinguishable from wild-type protein (Figure 1A). Catalytically inactive Aub localization to the pole plasm was also not perturbed (Figure 1A). Conversely, the localization of catalytically inactive Aub in nuage changed dramatically: Aub-CD formed large punctate granules in the cytoplasm more reminiscent of Ago3 granules than the smooth layer characteristic of wild-type Aub protein (Figure 2A; 2B). To further test if catalytically inactive Aub co-localizes with wild-type Ago3 we co-expressed the two proteins tagged with different fluorescent protein tags. Quantitative analysis shows that impairment of endonuclease activity causes

an increase in Aub co-localization with wild-type Ago3 (Pearson's correlation coefficient,  $R=0.897$  to  $0.910$  for Aub-CD and Ago3-WT vs.  $R=0.795$  to  $0.843$  for Aub-WT and Ago3-WT proteins;  $n = 8$  for all)(Supplementary Figure 1). Overall, impairment of Aub endonuclease activity leads to its localization in large punctate granules, where it co-localizes with Ago3.

The dynamics of catalytically inactive Aub and Ago3 measured by FRAP showed a decrease in the mobile fractions of both proteins, with an effect especially pronounced for Aub-CD: the mobile fraction of Aub-CD protein is  $31.9\% \pm 1.7\%$  compared to  $52.3\% \pm 2.0\%$  for wild-type Aub, and  $33.5\% \pm 2.4\%$  for Ago3-CD protein compared to  $40.9\% \pm 1.8\%$  for wild-type Ago3 (Figure 1F; Figure 2B). This result indicates that the endonuclease activity of Aub and Ago3 is required for their dissociation from nuage and catalytically impaired proteins appear trapped in granules. Importantly, impairment of Aub endonuclease activity leads to a decrease in mobility and leads to a localization pattern to granular structures indistinguishable from Ago3 granules.

#### ***Impairment of Aub endonuclease activity stabilizes its interaction with SpnE and Mael***

The formation of distinct granules and decreased mobility of catalytically impaired Aub might indicate stabilization of a normally transient complex prior to RNA target cleavage. Indeed, Aub-CD appeared to co-localize with other components of the piRNA pathway better than wild-type protein (Figure 3A; 3B). We tested interactions between wild-type and catalytically inactive Aub and other piRNA pathway proteins in *Drosophila* ovaries. Co-immunoprecipitations of tagged proteins in the presence of RNase revealed that Aub interacts with Qin and Krimp independently of its catalytic activity. In contrast, SpnE and Mael are present in a stable complex exclusively with Aub-CD and not the wild-type protein, while Interactions with Vasa protein or GFP were not detected (Figure 4A, 4B). We conclude that transient interactions between Aub with SpnE and Mael occur prior to cleavage of targeted RNA transcripts and these interactions are stabilized when cleavage cannot occur. Successful target cleavage disengages these interactions and promotes Aub mobility.

### ***Piwi devoid of piRNA forms a complex with Aub***

Mael was reported to enable Piwi co-transcriptional silencing functions in the nucleus [100]. Both Mael and SpnE also localize to the nucleus of germ cells while also localizing to perinuclear nuage (Figure 3). Alongside Mael and SpnE, we tested if Piwi might interact with Aub. Interestingly, the endonucleolytic mutant Aub-CD stability interacted with Piwi independently of RNA (Figure 5A). We next examined if piRNA binding deficient Piwi-YK also interacted with Aub (see Chapter 2). Surprisingly, not only could we detect Piwi-YK stably interacting with Aub-CD, but it also interacted with wildtype Aub (Figure 5B). The interaction between Piwi and Aub did not appear to require the Piwi N-terminal region since mutant Piwi protein lacking the first ~100 amino acids of the N-terminus co-purified with strongly with wildtype Aub (Figure 5C). In summary, our data indicate that Aub endonucleolytic function and piRNA loading of Piwi might be related, since Aub and Piwi can form a stable protein complex independent of RNA interactions.

## **DISCUSSION**

### ***piRNA directs Aub recruitment to nuage, while Aub nuclease activity releases it from nuage***

In Chapter 4 of this thesis, we present evidence that arginine methylation and the N-terminal region are not essential for the localization of Aub. Instead, our results from chapter three reveal an essential role of piRNA in Aub recruitment to nuage: an Aub point mutant that is not loaded with piRNA is completely delocalized from nuage and instead is diffused in the cytoplasm. This effect helps to explain the results of previous studies, which found that deficiencies in Rhino (Rhi), UAP56, and Cutoff (Cuff) proteins cause delocalization of Aub from nuage [63-65]. These three proteins are nuclear, precluding their direct interaction with Aub in the cytoplasm; however, as they are required for piRNA biogenesis, Aub remains unloaded and therefore mislocalizes in these mutants.

The strong dependence of Aub localization on piRNA binding suggests that Aub might be recruited to nuage through RNA-RNA interactions in the recognition of RNA

targets by guide piRNA cargo. Alternatively, binding of piRNA might lead to conformational changes in Aub that promote its interaction with other proteins in nuage (Chapter 4). Inactivation of the endonuclease activity of Aub has an effect that is opposite to the impairment of piRNA-binding: catalytically inactive Aub accumulates in large nuage granules and has lower mobility compared to the wild-type protein. These results indicate that the inability to cleave piRNA targets leads to accumulation of Aub in nuage. Together, these results point at piRNA-guided target recognition and cleavage as the main events that regulate Aub mobility in nuage.

***Aub interacts with piRNA pathway components, SpnE, and Mael in a pre-cleavage complex that recruits Piwi***

The inability of Aub to cleave piRNA targets stabilizes its interaction with two piRNA pathway proteins, SpnE and Mael. While neither SpnE nor Mael proteins co-purified with wild-type Aub, they are present in complex with Aub that is nuclease-deficient. This result suggests that SpnE and Mael transiently associate with Aub in a pre-cleavage complex. SpnE is a putative DExD-box RNA helicase that might help Aub access its RNA targets prior to cleavage. Mael has many functions in oogenesis, from participating in microtubule organizing center (MTOC) assembly to essential roles in the piRNA pathway [107, 151-153]. Mael was reported to act upstream of Piwi in the nucleus and both SpnE and Mael localize in the nucleus of germ cells (**Figure 3**)[100]. Cleavage deficient Aub also stably interacts with Piwi, which functions in co-transcriptional silencing of genomic loci expressing TE sequences in the nucleus of germ cells [101-103]. This finding is the first of its kind that directly links Piwi with the function of Aub in the cytoplasm. It remains to be understood how Mael, SpnE, and Piwi assemble with cleavage-impaired Aub, although it is possible that Mael and/or SpnE recruit Piwi to Aub prior to target cleavage.

The interaction detected between Aub and Piwi might indicate cross-talk between the cytoplasmic piRNA pathway and the nuclear Piwi pathway. In addition to cleavage-incompetent Aub stably interacting with Piwi protein, we found that piRNA-binding deficient Piwi also stably interacts with wild type Aub (Figure 5B). This interaction implies that at least a fraction of unloaded Piwi is normally in complex with Aub and that this interaction is responsive to Aub endonucleolytic function. Overall, our data indicate

that Aub scans RNAs exported from the nucleus for transcripts complementary to its associated piRNA. Recognition of target RNA by Aub-piRNA initiates complex assembly with other nuage components to recruit unloaded Piwi protein (Figure 6).

***The possibility that Aub directs the loading of piRNA into Piwi***

Could Aub be designating RNA substrates for loading into Piwi? To answer this question, the location of Piwi loading must be addressed. Since Aub is exclusive to the cytoplasm, we must address if Piwi could interact with Aub in the cytoplasm to load with piRNA. In Chapter 4, we show that an N-terminal domain swapped protein consisting of the Aub N-terminal region grafted on the PAZ-mid-PIWI domain of Piwi localizes to perinuclear nuage, but is loaded with piRNA most similar to Piwi piRNA (Chapter 4, Figure 2A, 2E). Additionally, a previous study reports that N-terminally truncated Piwi that is unable to localize to germ cell nuclei appears to be loaded with piRNA [130]. Lastly, in Chapter 2, we show that piRNA-binding deficient Piwi-YK protein is largely restricted from entering the nucleus (Chapter 2, Figure 2). Together these findings indicate that Piwi loading takes place in the cytoplasm before it is imported into the nucleus.

If Aub endonucleolytic activity initiated the loading of Piwi, an expectation would be that Aub piRNA and Piwi piRNA are phased. Phasing describes how multiple piRNA sequences generated from a single transcript would map in sequential intervals of periodic lengths of 30 nt or longer. Clear evidence of phasing between Aub/Ago3 piRNA and Piwi piRNA do not yet exist in literature, although personal communications with members of the Brennecke and Zamore groups do suggest that some instances of Piwi piRNA phasing with Aub and Ago3 piRNA do exist.

The hypothesis that Aub directs Piwi loading is further complicated by the paradox of piRNA strand orientation: Aub is loaded with primary piRNA generated from piRNA cluster transcripts and is predominantly anti-sense to TE transcripts. Likewise, Piwi piRNA consists of primary piRNA that is also anti-sense to TE transcripts. As such, there is no evidence based on sequencing data that would suggest that Piwi is loading with the products of Aub target cleavage, as is the case with Aub and Ago3 in the piRNA ping pong pathway (4P) discussed in Chapter 2.

Two putative mechanisms could explain how the paradox of piRNA strand orientation could be overcome: the first mechanism involves an unknown RNA-dependent RNA polymerase (RDRP) that uses the transcript targeted by Aub as a template to synthesize a nascent strand of RNA complementary to the target transcript. This strand would be of anti-sense orientation to the target transcript and could then be processed and loaded into Piwi. However, the existence of such an RDRP mechanism in the piRNA pathway of *Drosophila* has yet to be described [154, 155]. An alternative mechanism might be that Aub is loaded with an unprocessed piRNA precursor whose 3' end is only processed and loaded into Piwi upon successful target recognition by Aub. However, there is no evidence to support the existence of a piRNA intermediate longer than 30nt in wild type ovaries or in piRNA pathway mutant ovaries. While this data has exciting implications, more work is required to dissect whether these processes do in fact exist and how they function.

## REFERENCES

1. Slotkin, R. and R. Martienssen, *Transposable elements and the epigenetic regulation of the genome*. Nature reviews. Genetics, 2007. **8**(4): p. 272-285.
2. Siomi, M., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nature reviews. Molecular cell biology, 2011. **12**(4): p. 246-258.
3. Carmell, M., et al., *The Argonaute family: tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis*. Genes & development, 2002. **16**(21): p. 2733-2742.
4. Cox, D., et al., *A novel class of evolutionarily conserved genes defined by piwi are essential for stem cell self-renewal*. Genes & development, 1998. **12**(23): p. 3715-3727.
5. Kuramochi-Miyagawa, S., et al., *Mili, a mammalian member of piwi family gene, is essential for spermatogenesis*. Development (Cambridge, England), 2004. **131**(4): p. 839-849.
6. Kuramochi-Miyagawa, S., et al., *Two mouse piwi-related genes: miwi and mili*. Mechanisms of development, 2001. **108**(1-2): p. 121-133.
7. Lin, H. and A. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development (Cambridge, England), 1997. **124**(12): p. 2463-2476.
8. Houwing, S., et al., *A role for Piwi and piRNAs in germ cell maintenance and transposon silencing in Zebrafish*. Cell, 2007. **129**(1): p. 69-82.
9. Song, J.-J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science (New York, N.Y.), 2004. **305**(5689): p. 1434-1437.
10. Wang, Y., et al., *Structure of an argonaute silencing complex with a seed-containing guide DNA and target RNA duplex*. Nature, 2008. **456**(7224): p. 921-926.
11. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nature cell biology, 2009. **11**(5): p. 652-658.
12. Liu, K., et al., *Structural basis for recognition of arginine methylated Piwi proteins by the extended Tudor domain*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(43): p. 18398-18403.

13. Nishida, K., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. The EMBO journal, 2009. **28**(24): p. 3820-3831.
14. Vagin, V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes & development, 2009. **23**(15): p. 1749-1762.
15. Huang, H.-Y., et al., *Tdrd1 acts as a molecular scaffold for Piwi proteins and piRNA targets in zebrafish*. The EMBO journal, 2011. **30**(16): p. 3298-3308.
16. Mathioudakis, N., et al., *The multiple Tudor domain-containing protein TDRD1 is a molecular scaffold for mouse Piwi proteins and piRNA biogenesis factors*. RNA (New York, N.Y.), 2012. **18**(11): p. 2056-2072.
17. Aravin, A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Molecular cell, 2008. **31**(6): p. 785-799.
18. Aravin, A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS genetics, 2009. **5**(12).
19. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-1103.
20. Malone, C.D., et al., *Specialized piRNA pathways act in germline and somatic tissues of the Drosophila ovary*. Cell, 2009. **137**(3): p. 522-35.
21. Aravin, A., et al., *A novel class of small RNAs bind to MILI protein in mouse testes*. Nature, 2006. **442**(7099): p. 203-207.
22. Girard, A., et al., *A germline-specific class of small RNAs binds mammalian Piwi proteins*. Nature, 2006. **442**(7099): p. 199-202.
23. Grivna, S., et al., *A novel class of small RNAs in mouse spermatogenic cells*. Genes & development, 2006. **20**(13): p. 1709-1714.
24. Lau, N., et al., *Characterization of the piRNA complex from rat testes*. Science (New York, N.Y.), 2006. **313**(5785): p. 363-367.
25. Aravin, A.A., et al., *Developmentally Regulated piRNA Clusters Implicate MILI in Transposon Control*. Science, 2007. **316**(5825): p. 744-747.
26. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes & development, 2013. **27**(4): p. 390-399.



27. Sienski, G., D. Dönertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-980.
28. Ro, S., et al., *Cloning and expression profiling of testis-expressed piRNA-like RNAs*. RNA (New York, N.Y.), 2007. **13**(10): p. 1693-1702.
29. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Molecular cell, 2013. **50**(5): p. 749-761.
30. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Molecular cell, 2013. **50**(5): p. 762-777.
31. Muerdter, F., et al., *A genome-wide RNAi screen draws a genetic framework for transposon control and primary piRNA biogenesis in Drosophila*. Molecular cell, 2013. **50**(5): p. 736-748.
32. Olivieri, D., et al., *An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in Drosophila*. The EMBO journal, 2010. **29**(19): p. 3301-3317.
33. Ipsaro, J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-283.
34. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-287.
35. Voigt, F., et al., *Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc*. RNA (New York, N.Y.), 2012. **18**(12): p. 2128-2134.
36. Szakmary, A., et al., *The Yb protein defines a novel organelle and regulates male germline stem cell self-renewal in Drosophila melanogaster*. The Journal of Cell Biology, 2009. **185**.
37. Kawaoka, S., et al., *3' end formation of PIWI-interacting RNAs in vitro*. Molecular cell, 2011.
38. Gunawardane, L.S., et al., *A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in Drosophila*. Science, 2007. **315**(5818): p. 1587-90.

39. Huang, H., et al., *AGO3 Slicer activity regulates mitochondria-nuage localization of Armitage and piRNA amplification*. J Cell Biol, 2014. **206**(2): p. 217-30.
40. Li, C., et al., *Collapse of germline piRNAs in the absence of Argonaute3 reveals somatic piRNAs in flies*. Cell, 2009. **137**(3): p. 509-21.
41. Anand, A. and T. Kai, *The tudor domain protein kumo is required to assemble the nuage and to generate germline piRNAs in Drosophila*. EMBO J, 2012. **31**(4): p. 870-82.
42. Zhang, Z., et al., *Heterotypic piRNA Ping-Pong requires qin, a protein with both E3 ligase and Tudor domains*. Mol Cell, 2011. **44**(4): p. 572-84.
43. Xiol, J., et al., *RNA clamping by vasa assembles a piRNA amplifier complex on transposon transcripts*. Cell, 2014. **157**(7): p. 1698-711.
44. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-1299.
45. Klenov, M., et al., *Repeat-associated siRNAs cause chromatin silencing of retrotransposons in the Drosophila melanogaster germline*. Nucleic acids research, 2007. **35**(16): p. 5430-5438.
46. Chambeyron, S., et al., *piRNA-mediated nuclear accumulation of retrotransposon transcripts in the Drosophila female germline*. Proceedings of the National Academy of Sciences, 2008. **105**(39): p. 14964-14969.
47. Shpiz, S., et al., *rasiRNA pathway controls antisense expression of Drosophila telomeric retrotransposons in the nucleus*. Nucleic acids research, 2009. **37**(1): p. 268-278.
48. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic acids research, 2011. **39**(20): p. 8703-8711.
49. Rozhkov, N., M. Hammell, and G. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes & development, 2013. **27**(4): p. 400-412.
50. Soper, S., et al., *Mouse maelstrom, a component of nuage, is essential for spermatogenesis and transposon repression in meiosis*. Developmental cell, 2008. **15**(2): p. 285-297.
51. Song, S.U., et al., *Infection of the germ line by retroviral particles produced in the follicle cells: a possible mechanism for the mobilization of the gypsy retroelement of Drosophila*. Development, 1997. **124**(14): p. 2789-98.

52. al-Mukhtar, K.A. and A.C. Webb, *An ultrastructural study of primordial germ cells, oogonia and early oocytes in Xenopus laevis*. J Embryol Exp Morphol, 1971. **26**(2): p. 195-217.
53. Eddy, E.M. and S. Ito, *Fine structural and radioautographic observations on dense perinuclear cytoplasmic material in tadpole oocytes*. J Cell Biol, 1971. **49**(1): p. 90-108.
54. Mahowald, A.P., *Polar granules of Drosophila. II. Ultrastructural changes during early embryogenesis*. J Exp Zool, 1968. **167**(2): p. 237-61.
55. Hegner, R.W., *The History of the Germ Cells in the Paedogenetic Larva of Miastor*. Science, 1912. **36**(917): p. 124-6.
56. Hegner, R.W., *The Germ Cell Determinants in the Eggs of Chrysomelid Beetles*. Science, 1911. **33**(837): p. 71-2.
57. Hathaway, D.S. and G.G. Selman, *Certain aspects of cell lineage and morphogenesis studied in embryos of Drosophila melanogaster with an ultra-violet micro-beam*. J Embryol Exp Morphol, 1961. **9**: p. 310-25.
58. Okada, M., I.A. Kleinman, and H.A. Schneiderman, *Restoration of fertility in sterilized Drosophila eggs by transplantation of polar cytoplasm*. Dev Biol, 1974. **37**(1): p. 43-54.
59. Aravin, A. and D. Chan, *piRNAs meet mitochondria*. Developmental cell, 2011. **20**(3): p. 287-288.
60. Ipsaro, J.J., et al., *The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 279-83.
61. Czech, B., et al., *A transcriptome-wide RNAi screen in the Drosophila ovary reveals factors of the germline piRNA pathway*. Mol Cell, 2013. **50**(5): p. 749-61.
62. Ponting, C.P., *Tudor domains in proteins that interact with RNA*. Trends Biochem Sci, 1997. **22**(2): p. 51-2.
63. Klattenhoff, C., et al., *The Drosophila HPI homolog Rhino is required for transposon silencing and piRNA production by dual-strand clusters*. Cell, 2009. **138**(6): p. 1137-49.
64. Pane, A., et al., *The Cutoff protein regulates piRNA cluster expression and piRNA production in the Drosophila germline*. EMBO J, 2011. **30**(22): p. 4601-15.

65. Zhang, F., et al., *UAP56 couples piRNA clusters to the perinuclear transposon silencing machinery*. Cell, 2012. **151**(4): p. 871-84.
66. Lim, A.K. and T. Kai, *Unique germ-line organelle, nuage, functions to repress selfish genetic elements in Drosophila melanogaster*. Proc Natl Acad Sci U S A, 2007. **104**(16): p. 6714-9.
67. Patil, V.S. and T. Kai, *Repression of retroelements in Drosophila germline via piRNA pathway by the Tudor domain protein Tejas*. Curr Biol, 2010. **20**(8): p. 724-30.
68. Anne, J. and B.M. Mechler, *Valois, a component of the nuage and pole plasm, is involved in assembly of these structures, and binds to Tudor and the methyltransferase Capsuleen*. Development, 2005. **132**(9): p. 2167-77.
69. Kirino, Y., et al., *Arginine methylation of Aubergine mediates Tudor binding and germ plasm localization*. RNA, 2010. **16**(1): p. 70-8.
70. Liu, H., et al., *Structural basis for methylarginine-dependent recognition of Aubergine by Tudor*. Genes Dev, 2010. **24**(17): p. 1876-81.
71. Kirino, Y., et al., *Arginine methylation of vasa protein is conserved across phyla*. J Biol Chem, 2010. **285**(11): p. 8148-54.
72. Liu, L., et al., *PAPI, a novel TUDOR-domain protein, complexes with AGO3, ME31B and TRAL in the nuage to silence transposition*. Development, 2011. **138**(9): p. 1863-73.
73. Hay, B., et al., *Identification of a component of Drosophila polar granules*. Development, 1988. **103**(4): p. 625-40.
74. Anne, J., *Targeting and anchoring Tudor in the pole plasm of the Drosophila oocyte*. PLoS One, 2010. **5**(12): p. e14362.
75. Arkov, A.L., et al., *The role of Tudor domains in germline development and polar granule architecture*. Development, 2006. **133**(20): p. 4053-62.
76. Anne, J., et al., *Arginine methyltransferase Capsuleen is essential for methylation of spliceosomal Sm proteins and germ cell formation in Drosophila*. Development, 2007. **134**(1): p. 137-46.
77. Thomson, T. and P. Lasko, *Drosophila tudor is essential for polar granule assembly and pole cell specification, but not for posterior patterning*. Genesis, 2004. **40**(3): p. 164-70.

78. Kugler, J.M. and P. Lasko, *Localization, anchoring and translational control of oskar, gurken, bicoid and nanos mRNA during Drosophila oogenesis*. Fly (Austin), 2009. **3**(1): p. 15-28.
79. Liang, L., W. Diehl-Jones, and P. Lasko, *Localization of vasa protein to the Drosophila pole plasm is independent of its RNA-binding and helicase activities*. Development, 1994. **120**(5): p. 1201-11.
80. Megosh, H.B., et al., *The role of PIWI and the miRNA machinery in Drosophila germline determination*. Curr Biol, 2006. **16**(19): p. 1884-94.
81. Aravin, A.A., et al., *Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the D. melanogaster germline*. Curr Biol, 2001. **11**(13): p. 1017-27.
82. Bozzetti, M.P., et al., *The Ste locus, a component of the parasitic cry-Ste system of Drosophila melanogaster, encodes a protein that forms crystals in primary spermatocytes and mimics properties of the beta subunit of casein kinase 2*. Proc Natl Acad Sci U S A, 1995. **92**(13): p. 6067-71.
83. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-760.
84. Vagin, V.V., et al., *A distinct small RNA pathway silences selfish genetic elements in the germline*. Science, 2006. **313**(5785): p. 320-4.
85. Aravin, A.A., et al., *Dissection of a natural RNA silencing process in the Drosophila melanogaster germ line*. Mol Cell Biol, 2004. **24**(15): p. 6742-50.
86. Handler, D., et al., *A systematic analysis of Drosophila TUDOR domain-containing proteins identifies Vreteno and the Tdrd12 family as essential primary piRNA pathway factors*. EMBO J, 2011. **30**(19): p. 3977-93.
87. Stapleton, W., S. Das, and B.D. McKee, *A role of the Drosophila homeless gene in repression of Stellate in male meiosis*. Chromosoma, 2001. **110**(3): p. 228-40.
88. Kibanov, M.V., et al., *A novel organelle, the piNG-body, in the nuage of Drosophila male germ cells is associated with piRNA-mediated gene silencing*. Mol Biol Cell, 2011. **22**(18): p. 3410-9.

89. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(46): p. 18760-5.
90. Saito, K., et al., *A regulatory circuit for piwi by the large Maf gene traffic jam in Drosophila*. Nature, 2009. **461**(7268): p. 1296-9.
91. Aravin, A.A., G.J. Hannon, and J. Brennecke, *The Piwi-piRNA pathway provides an adaptive defense in the transposon arms race*. Science, 2007. **318**(5851): p. 761-4.
92. Siomi, M.C., et al., *PIWI-interacting small RNAs: the vanguard of genome defence*. Nat Rev Mol Cell Biol, 2011. **12**(4): p. 246-58.
93. Ma, J.B., K. Ye, and D.J. Patel, *Structural basis for overhang-specific small interfering RNA recognition by the PAZ domain*. Nature, 2004. **429**(6989): p. 318-22.
94. Ma, J.B., et al., *Structural basis for 5'-end-specific recognition of guide RNA by the A. fulgidus Piwi protein*. Nature, 2005. **434**(7033): p. 666-70.
95. Yuan, Y.R., et al., *Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage*. Mol Cell, 2005. **19**(3): p. 405-19.
96. Song, J.J., et al., *Crystal structure of Argonaute and its implications for RISC slicer activity*. Science, 2004. **305**(5689): p. 1434-7.
97. Lin, H. and A.C. Spradling, *A novel group of pumilio mutations affects the asymmetric division of germline stem cells in the Drosophila ovary*. Development, 1997. **124**(12): p. 2463-76.
98. Schmidt, A., et al., *Genetic and molecular characterization of sting, a gene involved in crystal formation and meiotic drive in the male germ line of Drosophila melanogaster*. Genetics, 1999. **151**(2): p. 749-60.
99. Brennecke, J., et al., *Discrete small RNA-generating loci as master regulators of transposon activity in Drosophila*. Cell, 2007. **128**(6): p. 1089-103.
100. Sienski, G., D. Donertas, and J. Brennecke, *Transcriptional silencing of transposons by Piwi and maelstrom and its impact on chromatin state and gene expression*. Cell, 2012. **151**(5): p. 964-80.

101. Le Thomas, A., et al., *Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state*. Genes Dev, 2013. **27**(4): p. 390-9.
102. Rozhkov, N.V., M. Hammell, and G.J. Hannon, *Multiple roles for Piwi in silencing Drosophila transposons*. Genes Dev, 2013. **27**(4): p. 400-12.
103. Shpiz, S., et al., *Mechanism of the piRNA-mediated silencing of Drosophila telomeric retrotransposons*. Nucleic Acids Res, 2011. **39**(20): p. 8703-11.
104. Eddy, E.M., *Germ plasm and the differentiation of the germ cell line*. Int Rev Cytol, 1975. **43**: p. 229-80.
105. Mahowald, A.P., *Polar granules of drosophila. IV. Cytochemical studies showing loss of RNA from polar granules during early stages of embryogenesis*. J Exp Zool, 1971. **176**(3): p. 345-52.
106. Patil, V.S. and T. Kai, *Repression of Retroelements in Drosophila Germline via piRNA Pathway by the Tudor Domain Protein Tejas*. Curr Biol, 2010.
107. Findley, S.D., et al., *Maelstrom, a Drosophila spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage*. Development, 2003. **130**(5): p. 859-71.
108. Lim, A.K., L. Tao, and T. Kai, *piRNAs mediate posttranscriptional retroelement silencing and localization to pi-bodies in the Drosophila germline*. J Cell Biol, 2009. **186**(3): p. 333-42.
109. Nishida, K.M., et al., *Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in Drosophila germlines*. EMBO J, 2009. **28**(24): p. 3820-31.
110. Schupbach, T. and E. Wieschaus, *Female sterile mutations on the second chromosome of Drosophila melanogaster. II. Mutations blocking oogenesis or altering egg morphology*. Genetics, 1991. **129**(4): p. 1119-36.
111. Harris, A.N. and P.M. Macdonald, *Aubergine encodes a Drosophila polar granule component required for pole cell formation and related to eIF2C*. Development, 2001. **128**(14): p. 2823-32.
112. Bossing, T., C.S. Barros, and A.H. Brand, *Rapid tissue-specific expression assay in living embryos*. Genesis, 2002. **34**(1-2): p. 123-6.
113. Djuranovic, S., et al., *Allosteric regulation of Argonaute proteins by miRNAs*. Nat Struct Mol Biol, 2010. **17**(2): p. 144-50.

114. Parker, J.S., S.M. Roe, and D. Barford, *Structural insights into mRNA recognition from a PIWI domain-siRNA guide complex*. Nature, 2005. **434**(7033): p. 663-6.
115. Nishimasu, H., et al., *Structure and function of Zucchini endoribonuclease in piRNA biogenesis*. Nature, 2012. **491**(7423): p. 284-7.
116. Snee, M.J. and P.M. Macdonald, *Live imaging of nuage and polar granules: evidence against a precursor-product relationship and a novel role for Oskar in stabilization of polar granule components*. J Cell Sci, 2004. **117**(Pt 10): p. 2109-20.
117. Nagao, A., et al., *Gender-Specific Hierarchy in Nuage Localization of PIWI-Interacting RNA Factors in Drosophila*. Front Genet, 2011. **2**: p. 55.
118. Wang, Y., et al., *Nucleation, propagation and cleavage of target RNAs in Ago silencing complexes*. Nature, 2009. **461**(7265): p. 754-61.
119. Aravin, A.A., et al., *Cytoplasmic compartmentalization of the fetal piRNA pathway in mice*. PLoS Genet, 2009. **5**(12): p. e1000764.
120. Aravin, A.A., et al., *A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice*. Mol Cell, 2008. **31**(6): p. 785-99.
121. Shoji, M., et al., *The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline*. Dev Cell, 2009. **17**(6): p. 775-87.
122. Handler, D., et al., *The genetic makeup of the Drosophila piRNA pathway*. Mol Cell, 2013. **50**(5): p. 762-77.
123. Kirino, Y., et al., *Arginine methylation of Piwi proteins catalysed by dPRMT5 is required for Ago3 and Aub stability*. Nat Cell Biol, 2009. **11**(5): p. 652-8.
124. Xiol, J., et al., *A role for Fkbp6 and the chaperone machinery in piRNA amplification and transposon silencing*. Mol Cell, 2012. **47**(6): p. 970-9.
125. Cora, E., et al., *The MID-PIWI module of Piwi proteins specifies nucleotide- and strand-biases of piRNAs*. RNA, 2014. **20**(6): p. 773-81.
126. Vermaak, D., S. Henikoff, and H.S. Malik, *Positive selection drives the evolution of rhino, a member of the heterochromatin protein 1 family in Drosophila*. PLoS Genet, 2005. **1**(1): p. 96-108.



127. Obbard, D.J., et al., *The evolution of RNAi as a defence against viruses and transposable elements*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1513): p. 99-115.
128. Simkin, A., et al., *Recurrent and recent selective sweeps in the piRNA pathway*. Evolution, 2013. **67**(4): p. 1081-90.
129. Comeron, J.M., *A method for estimating the numbers of synonymous and nonsynonymous substitutions per site*. J Mol Evol, 1995. **41**(6): p. 1152-9.
130. Klenov, M.S., et al., *Separation of stem cell maintenance and transposon silencing functions of Piwi protein*. Proc Natl Acad Sci U S A, 2011. **108**(46): p. 18760-5.
131. Anne, J., *C-terminal moiety of Tudor contains its in vivo activity in Drosophila*. PLoS One, 2010. **5**(12): p. e14378.
132. Creed, T.M., et al., *Novel role of specific Tudor domains in Tudor-Aubergine protein complex assembly and distribution during Drosophila oogenesis*. Biochem Biophys Res Commun, 2010. **402**(2): p. 384-9.
133. Vagin, V.V., G.J. Hannon, and A.A. Aravin, *Arginine methylation as a molecular signature of the Piwi small RNA pathway*. Cell Cycle, 2009. **8**(24): p. 4003-4.
134. Arkov, A.L. and A. Ramos, *Building RNA-protein granules: insight from the germline*. Trends Cell Biol, 2010. **20**(8): p. 482-90.
135. Siomi, M.C., T. Mannen, and H. Siomi, *How does the royal family of Tudor rule the PIWI-interacting RNA pathway?* Genes Dev, 2010. **24**(7): p. 636-46.
136. Cavey, M., et al., *Drosophila valois encodes a divergent WD protein that is required for Vasa localization and Oskar protein accumulation*. Development, 2005. **132**(3): p. 459-68.
137. Ghabrial, A. and T. Schupbach, *Activation of a meiotic checkpoint regulates translation of Gurken during Drosophila oogenesis*. Nat Cell Biol, 1999. **1**(6): p. 354-7.
138. Klattenhoff, C., et al., *Drosophila rasiRNA pathway mutations disrupt embryonic axis specification through activation of an ATR/Chk2 DNA damage response*. Dev Cell, 2007. **12**(1): p. 45-55.

139. Vagin, V.V., et al., *Proteomic analysis of murine Piwi proteins reveals a role for arginine methylation in specifying interaction with Tudor family members*. Genes Dev, 2009. **23**(15): p. 1749-62.
140. Bedford, M.T. and S.G. Clarke, *Protein arginine methylation in mammals: who, what, and why*. Mol Cell, 2009. **33**(1): p. 1-13.
141. Krause, C.D., et al., *Protein arginine methyltransferases: evolution and assessment of their pharmacological and therapeutic potential*. Pharmacol Ther, 2007. **113**(1): p. 50-87.
142. Wang, J., et al., *Mili interacts with tudor domain-containing protein 1 in regulating spermatogenesis*. Curr Biol, 2009. **19**(8): p. 640-4.
143. Saxe, J.P., et al., *Tdrkh is essential for spermatogenesis and participates in primary piRNA biogenesis in the germline*. EMBO J, 2013. **32**(13): p. 1869-85.
144. Pandey, R.R., et al., *Tudor domain containing 12 (TDRD12) is essential for secondary PIWI interacting RNA biogenesis in mice*. Proc Natl Acad Sci U S A, 2013. **110**(41): p. 16492-7.
145. Lau, N.C., et al., *An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans*. Science, 2001. **294**(5543): p. 858-62.
146. Comeron, J.M., *K-Estimator: calculation of the number of nucleotide substitutions per site and the confidence intervals*. Bioinformatics, 1999. **15**(9): p. 763-4.
147. Lau, N.C., et al., *Abundant primary piRNAs, endo-siRNAs, and microRNAs in a Drosophila ovary cell line*. Genome Res, 2009. **19**(10): p. 1776-85.
148. O'Donnell, K.A. and J.D. Boeke, *Mighty Piwis defend the germline against genome intruders*. Cell, 2007. **129**(1): p. 37-44.
149. Malone, C.D. and G.J. Hannon, *Molecular evolution of piRNA and transposon control pathways in Drosophila*. Cold Spring Harb Symp Quant Biol, 2009. **74**: p. 225-34.
150. Gillespie, D.E. and C.A. Berg, *Homeless is required for RNA localization in Drosophila oogenesis and encodes a new member of the DE-H family of RNA-dependent ATPases*. Genes Dev, 1995. **9**(20): p. 2495-508.
151. Clegg, N.J., et al., *Maelstrom is required to position the MTOC in stage 2-6 Drosophila oocytes*. Dev Genes Evol, 2001. **211**(1): p. 44-8.

152. Clegg, N.J., et al., *maelstrom is required for an early step in the establishment of Drosophila oocyte polarity: posterior localization of grk mRNA*. Development, 1997. **124**(22): p. 4661-71.
153. Theurkauf, W.E., et al., *Reorganization of the cytoskeleton during Drosophila oogenesis: implications for axis specification and intercellular transport*. Development, 1992. **115**(4): p. 923-36.
154. Lipardi, C. and B.M. Paterson, *Identification of an RNA-dependent RNA polymerase in Drosophila involved in RNAi and transposon suppression*. Proc Natl Acad Sci U S A, 2009. **106**(37): p. 15645-50.
155. Lipardi, C. and B.M. Paterson, *Retraction for Lipardi and Paterson, "Identification of an RNA-dependent RNA polymerase in Drosophila involved in RNAi and transposon suppression"*. Proc Natl Acad Sci U S A, 2011. **108**(36): p. 15010.

## **MATERIALS AND METHODS**

### *Generation of Transgenic Fly Lines*

Transgenic protein constructs for injection were generated using Gateway cloning system (Life Technologies). Genes were obtained by PCR from *Drosophila melanogaster* template cDNA taken from ovaries or testes of adult flies. Point mutations were engineered by overlap PCR and inserted in the pENTR-D-TOPO vector. Transgenes were cloned into the pUASP-Gateway-phiC31 fly injection vector derived from pCasPeR5-phiC31 vector containing GFP, mKate2 or Strep-FLAG tags. All transgenes created for this manuscript contain N-terminal tags, with the exception of SpnE and Zuc transgenes, which were C-terminally tagged. The expression of each transgene was controlled using the yeast upstream activation sequence promoter (UASp) coupled with a maternal  $\alpha$ -Tubulin67c-Gal4-VP16 (M $\alpha$ G4) driver, which showed strong expression in nurse cells starting at stages 2-4 of oogenesis onwards to the oocyte. Transgenes were generated in flies by PhiC31-mediated transformation (BestGene) using PhiC31 landing pads on either chromosome two (BDSC #9736) or chromosome three (BDSC #9750).

### *Co-immunoprecipitation, & Western Blots:*

For co-immunoprecipitation of proteins from fly ovaries, 50 pairs of ovaries from flies fed yeast were lysed in 300 $\mu$ L NT2 buffer (50mM Tris at pH 7.4, 150 NaCl, 1mM MgCl<sub>2</sub>, 0.05% Igepal, EDTA-free Complete Protease Inhibitor Cocktail [Roche]). Supernatant was cleared by centrifugation at 4,000 x g for 10 minutes at 4°C. Input sample was collected from the supernatant, followed by splitting the sample into two aliquots treated with or without 100 $\mu$ g/mL RNase A for 10 minutes at 25°C. Anti-FLAG M2 beads (Sigma Aldrich) were blocked in 5 mg/ml BSA for 10 minutes at 4°C, followed by washing in NT2 buffer. Beads were added to the supernatant and rotated at 4°C for 90min. Unbound fraction was collected for analysis and beads were washed 3 times 5min. in 300 $\mu$ L NT2 buffer, and eluted by boiling in reducing SDS loading buffer, followed by running one third of the IP material on an 8% SDS-PAAG. Western Blots were probed with rabbit anti-GFP (Covance) or anti-FLAG M2 (Sigma Aldrich) antibody at 1:3,000 concentration. Anti-Vas antibody was generously provided by Dr. Paul Lasko and was used at 1:20,000 concentration.

## *Microscopy*

Ovaries were fixed in 4% PFA in PBS for 20 minutes, permeabilized in 1% Triton-X100 in PBS, and DAPI stained (Sigma-Aldrich). Actin was stained in specimen as indicated in the figure legends with Rhodamine Phalloidin (Life Technologies) according to the manufacturer's protocol. Ovaries were washed in PBS and mounted in Vectashield medium (Vector Labs). S2 cells were allowed to settle on coverslips treated with Poly-L-Lysine (Sigma-Aldrich). After gentle washing, cells were fixed in 0.5% PFA in PBS for 20 minutes followed by staining with DAPI (Sigma-Aldrich), washed, and mounted in Vectashield medium (Vector Labs). Images were captured using an Axiolmager microscope; an Apotome structured illumination system was used for optical sections (Carl Zeiss).

## *Fluorescence Recovery After Photobleaching (FRAP)*

For each construct, at least 20 independent FRAP experiments using ovaries expressing a single GFP-tagged transgene under the MaG4 driver were performed. FRAP experiments were captured on a Zeiss LSM710 confocal microscope (Carl Zeiss AIM) equipped with a 25x/0.8 NA Imm Corr multi-immersion objective lens and operating Zeiss Zen Black software. Image acquisition for all experiments utilized an identical 488nm AOTF laser power setting of 7% to ensure laser power was not influencing measurements. PMT gain settings were variably set to accommodate the expression level of each GFP-tagged protein. Images were acquired at 256x256 pixel resolution at 0.07 $\mu$ m pixel size and scan speed of 614.4ms per frame with 1.0  $\mu$ s pixel dwell time. A single bleach region was defined for each experiment, consisting of a region of 7x7 pixels equal to 0.49  $\mu$ m x 0.49  $\mu$ m and was bleached by a single iteration of 100% laser power from 488, 561, and 633nm wavelengths. Five initial pre-bleach images were captured prior to bleaching and 115 subsequent post-bleach images were acquired every 614.4ms to assess fluorescence recovery in the bleach zone. FIJI Is Just ImageJ (FIJI; <http://fiji.sc/>) software was used to analyze FRAP experiments. To account for background and photo-bleaching effects during acquisition, the mean intensity values from the bleach zone (*BL*), the background zone (*BG*), and the reference signal zone (*REF*) were used to calculate the corrected BL (*BL<sub>corr</sub>*) for each acquisition frame using the equation:

$$BL_{corr} = \frac{[BL - BG]}{[REF - BG]}$$

$BL_{corr}$  values were normalized to the mean of five pre-bleach values, which were used to estimate 100% fluorescence intensity. Using the curve fitter module of FIJI, the normalized post-bleach data was fit to an exponential recovery model:

$$y = a(1 - e^{-bx}) + c$$

The mobile fraction was obtained by the sum of coefficients  $a$  and  $c$ , which describes the maximal extent of recovery for each experiment. The mean mobile fraction for each line was calculated from at least 20 replicate FRAP experiments taken from at least four animals from the same stable line.

#### *Nuage: Cytoplasm localization analysis*

The occupancy of GFP-tagged protein in either nuage versus cytoplasm was calculated by obtaining the ratio of mean signal intensities of a perinuclear area ( $N$ ), an adjacent cytoplasmic area ( $C$ ) of equal size, and a background area ( $G$ ) using the following equation:

$$N:C = \frac{[N - G]}{[C - G]}$$

At least 80 different imaged nurse cells were analyzed to obtain our measurements. All images were acquired from fixed specimen using the Apotome structured illumination system and a 40x oil-immersion objective.

#### *Co-localization analysis*

Images of egg chambers co-expressing one mK2-tagged protein and one GFP-tagged protein were acquired from fixed specimen using the Apotome structured illumination system and a 40x oil-immersion objective. All measurements of intensity in this manuscript utilized FIJI and built-in functions of the software. The “Coloc2” plug-in for

FIJI was used to obtain the Pearson's correlation coefficient (R) for each protein pair. Eight independent nuclei were analyzed per protein pair and a mean R value was obtained by converting values using a Fisher's R to Z transformation calculator (<http://vassarstats.net/tabs.html#fisher>). Mean and standard error values for each protein set were then calculated and converted back to R-space using a reverse calculator ([http://onlinestatbook.com/calculators/fisher\\_z.html](http://onlinestatbook.com/calculators/fisher_z.html)).

### *Semi-Quantitative piRNA Binding Affinity Analysis*

Normalized Signal Intensities of piRNA levels after IP of GFP-tagged Aub and Ago3 transgenes was achieved by a normalized semi-quantitative approach. The gel band intensity corresponding to 29-19nt piRNA size range from autoradiographs ( $I_{piRNA}$ ) were normalized to western blot band intensities corresponding to GFP-tagged transgenes after IP ( $I_{protein}$ ). To account for potential losses in piRNA yield due to RNA purification, IP were spiked with 5pmol of synthesized 42-nt (42M) RNA oligomer prior to proteinase K digestion and phenol extracton. Autoradiographed gel band intensities from the 42M oligomer ( $I_{42M}$ ) were subtracted from western blot band intensities. Background signal intensities from adjacent lane regions of equal area were subtracted from each band intensity measurement. These values were then used to obtain the piRNA normalized signal intensity using the following calculation:

$$piRNA \text{ normalized signal intensity} = \frac{[I_{piRNA} - I_{BGpiRNA}]}{[I_{protein} - I_{BGprotein}] - [I_{42M} - I_{BG42M}]}$$

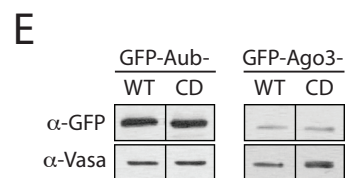
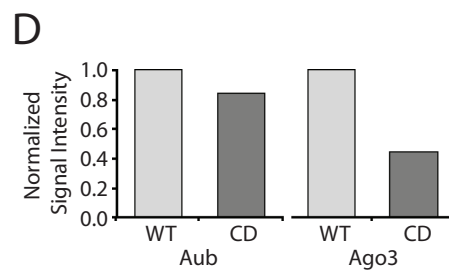
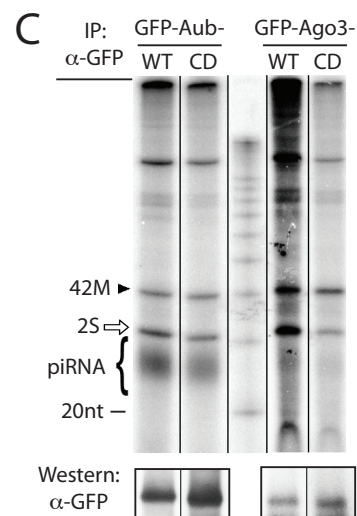
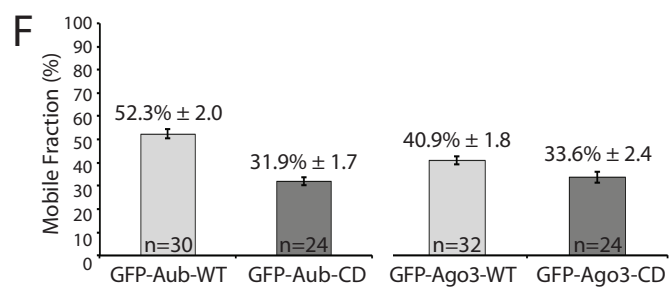
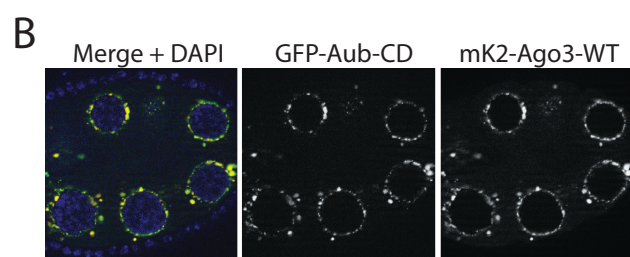
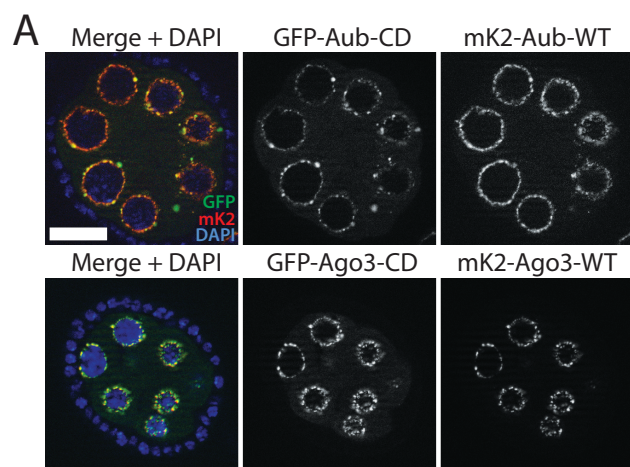
Band intensities were obtained using FIJI software and values graphed using Microsoft Excel.

## FIGURES AND FIGURE LEGENDS

***Figure 1. Endonuclease activity is required for the mobility of Aub and Ago3 in nuage.***

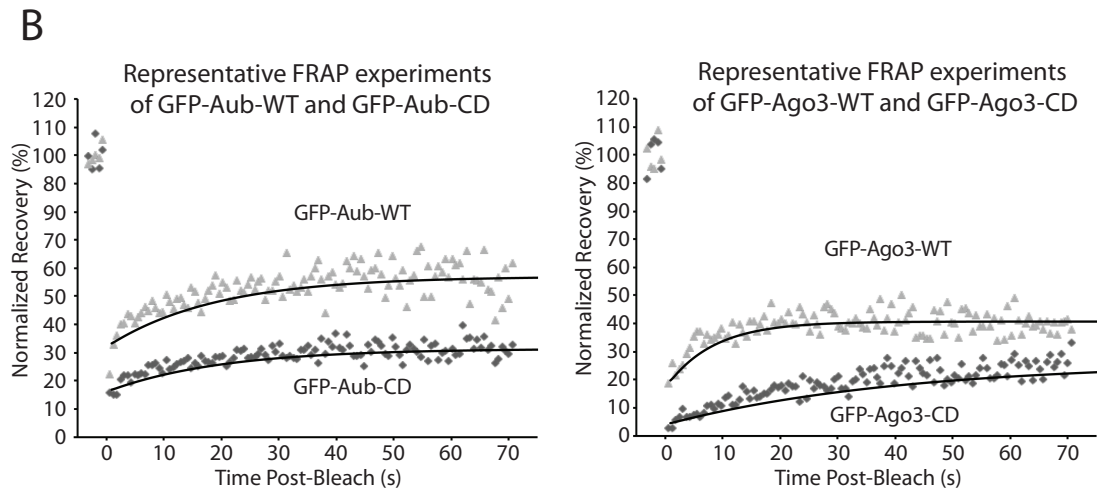
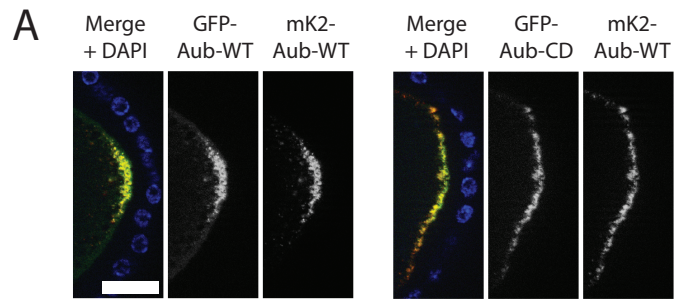
(A) The localizations of catalytically inactive Aub-CD (green) and wild-type (red) Aub proteins in nuage are not identical (upper panel). Aub-CD has a more granular localization compared to wild-type protein. Ago3-CD (green) localization appears similar to the wild-type Ago3 (red; lower panel). DNA is stained with DAPI (blue). Scale bar: 20µm. (B) Catalytically dead Aub-CD (green) strongly co-localizes with wild-type Ago3 (red). DNA is stained with DAPI (blue). (C) Catalytically inactive Aub and Ago3 co-purify with reduced amount of piRNA. piRNAs were isolated from immunoprecipitates of wild-type and catalytically dead Aub and Ago3. The strong band at 30 nt is an abundant 2S ribosomal RNA (rRNA) (white arrow). A 42nt RNA (42M) was spiked into each immunoprecipitate to account for loss of RNA during isolation and efficiency of RNA labeling (arrowhead). The levels of each protein after immune-purification are shown by western blot (lower panels). An equal fraction of immuno-purified material was loaded in each lane. (D) Quantification of piRNA loading of wild-type and mutant Aub and Ago3 shown on panel (C). The signal of radiolabeled piRNA was normalized to the amount of immunoprecipitated GFP-tagged Aub or Ago3 protein and to the amount of 42nt spike RNA (42M), which was used to normalize against losses during RNA purification. (E) The levels of Aub-CD and Ago3-CD are similar to respective wild-type proteins as measured by Western blot of ovary lysates. The amount of Ago3-CD is reduced 25% compared to wild-type proteins when normalized to the germ cell specific protein Vas. All transgenes are expressed using the same driver. (F) The mobility of catalytically dead Aub and Ago3 proteins is decreased. The fraction of mobile Aub-CD and Ago3-CD from replicate FRAP experiments ( $n=24$  each) are reduced by approximately 20% and 7%, respectively, compared to wild-type proteins. Bars show standard error values for each set of experiments.





***Figure 2. Catalytic activity is required for the mobility of Aub and Ago3, but does not affect pole plasm localization or loading of piRNA into Aub.***

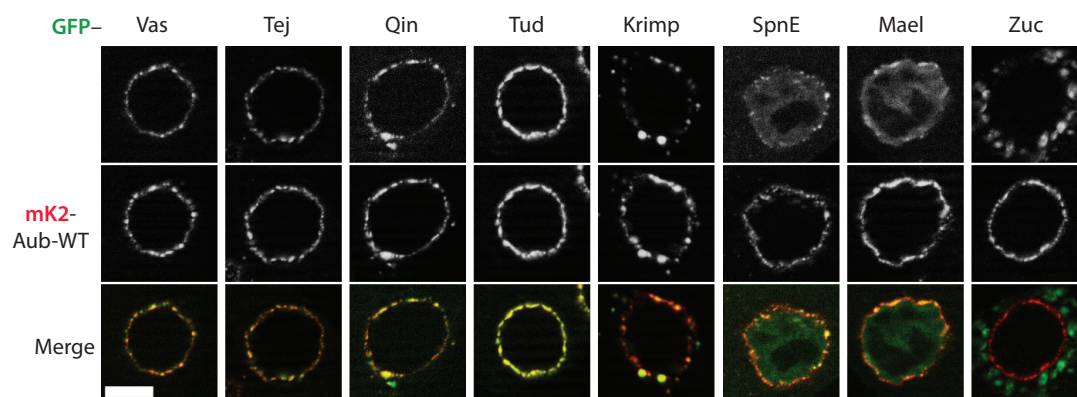
(A) The pole plasm localization of GFP-Aub-CD (green) is similar to mK2-Aub-WT (red) when co-expressed in the same ovary (right panel). DNA is stained with DAPI (blue). Scale Bar: 20 $\mu$ m. (B) Left panel: Representative FRAP experiments show that the normalized recovery of GFP-Aub-CD is approximately 20% lower compared to GFP-Aub-WT. Right panel: Representative FRAP experiments show that the normalized recovery of GFP-Ago3-CD is approximately 7% lower compared to GFP-Ago3-WT. The mobile fraction was determined by modeling the recovery to an exponential recovery curve (superimposed on respective data-sets).



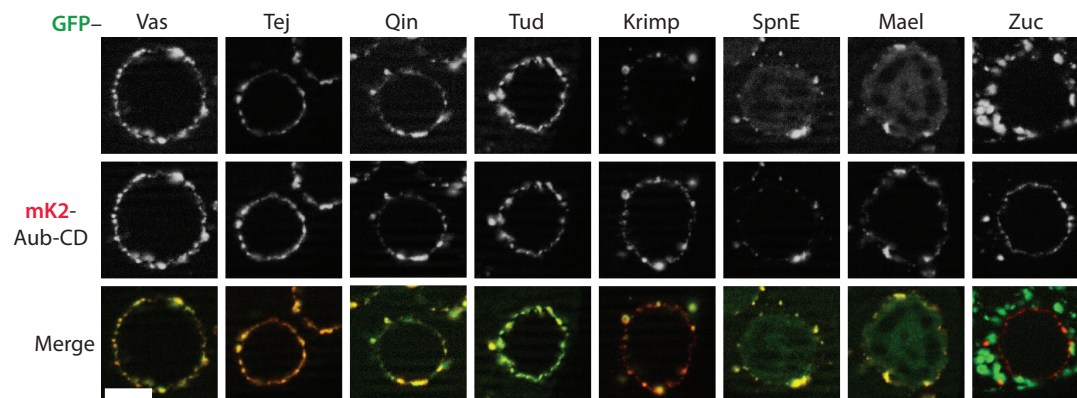
***Figure 3. Co-localization of catalytically dead Aub (Aub-CD) with other nuage components.***

(A) Wildtype and (B) endonucleolytic mutant Aub-CD (mK2-Aub-CD, red) was co-expressed with GFP-tagged piRNA pathway components (green) using the same driver. Vas, Tej, Qin, Tud, and Krimp show strong co-localization Aub-CD in nuage granules. SpnE and Mael localize in nuage but also in the nucleus. Zuc shows little to no co-localization with Aub-CD, since it is associated with mitochondria outside of the perinuclear space. Scale bar: 10µm.

A

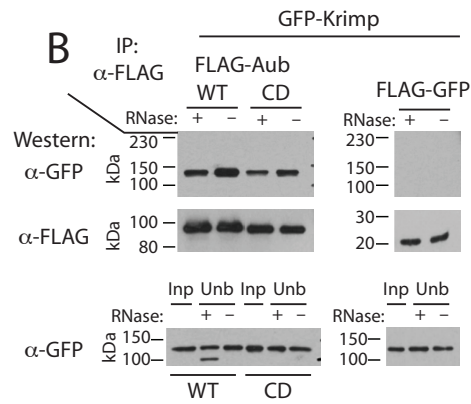
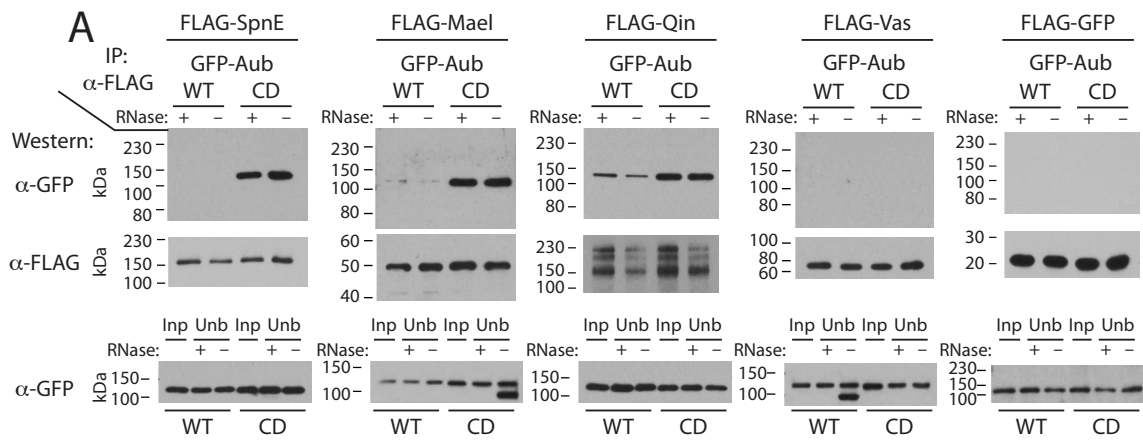


B



***Figure 4. Catalytic inactivation of Aub leads to stabilization of its interactions with SpnE and Mael.***

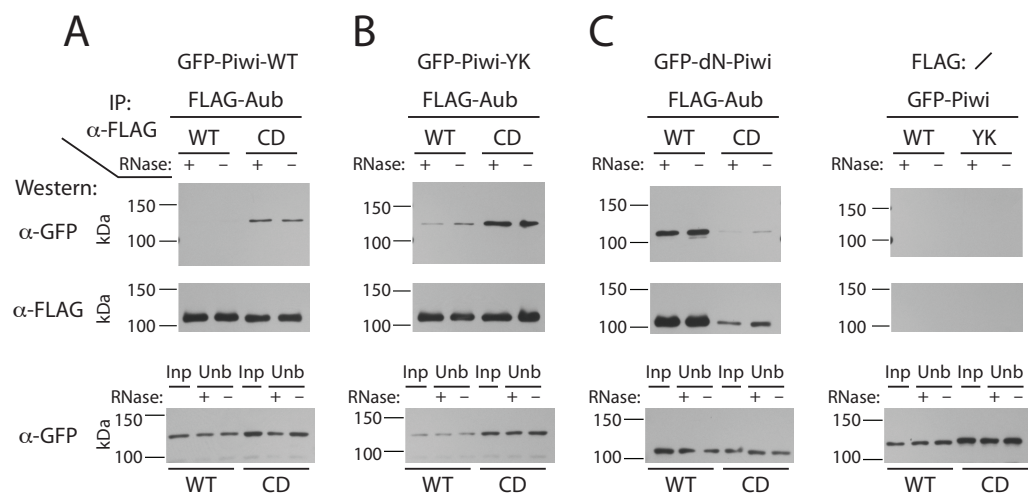
(A) Catalytically dead Aub (Aub-CD) forms a stable complex with Spn-E and Mael proteins independently of RNA that is not detected in wild-type Aub (Aub-WT). Impairment of endonuclease activity of Aub also enhances its interaction with Qin. Interactions with Vas or FLAG-tagged GFP were not detected. (B) Krimp protein forms a complex with Aub protein independently of Aub endonuclease activity in *Drosophila* ovary. Immunoprecipitations were performed from ovaries of transgenic flies that express FLAG- and GFP-tagged proteins in germ cells under control of the MaG4 driver. FLAG-tagged proteins were immunoprecipitated in the presence (+) or absence (-) of RNase A followed by Western blotting to detect GFP- and FLAG-tagged proteins. The lower panel shows Input (Inp) and the unbound fractions (Unb) collected after immunoprecipitation.



***Figure 5. The interaction between Aub and Piwi is stabilized by nuclease-deficiency of Aub and piRNA-binding deficiency of Piwi.***

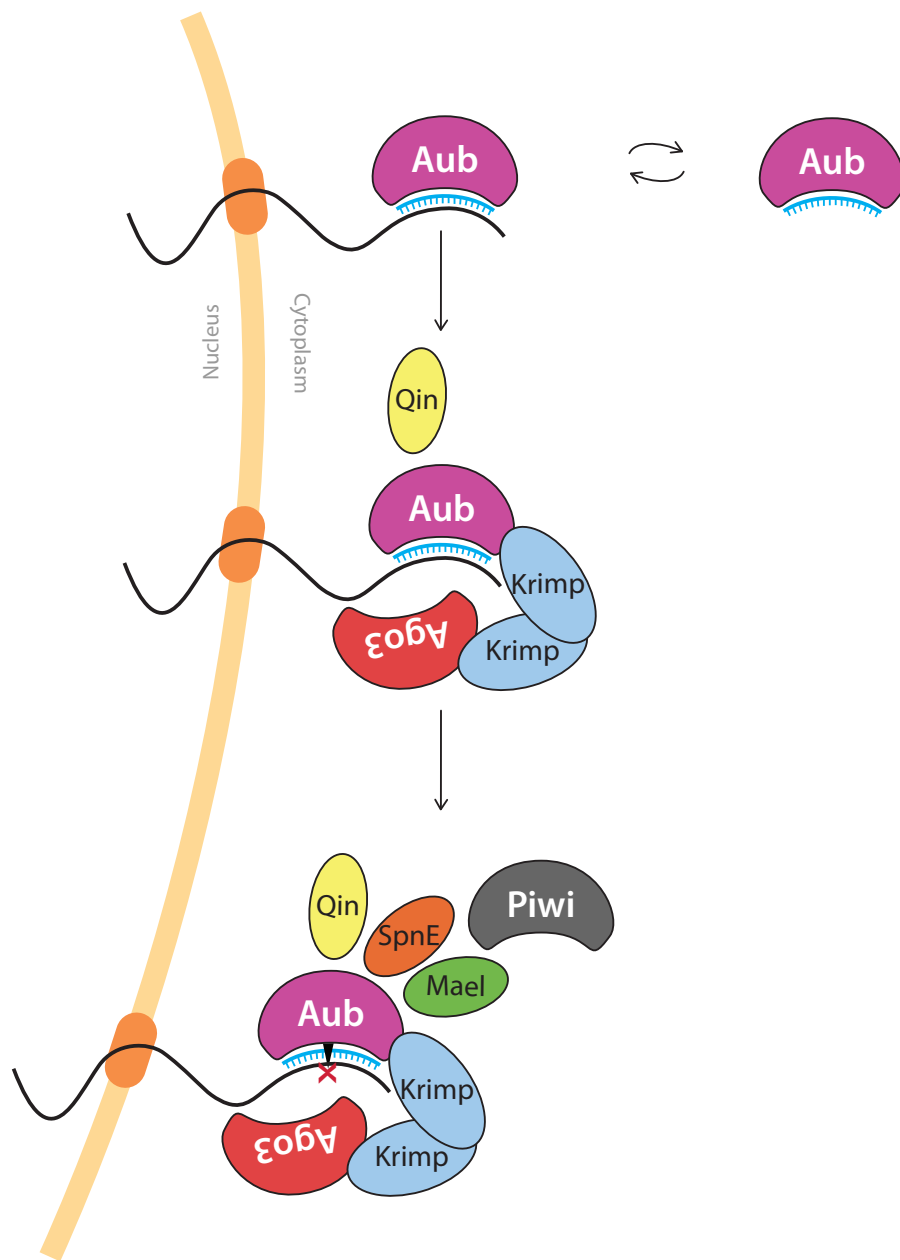
(A) Endonucleolytic mutant Aub-CD stably interacts with wildtype Piwi protein. (B) Wildtype Aub stably interacts with piRNA-binding deficient Piwi-YK, while the interaction is enhanced by inactivation of Aub endonucleolytic activity. (C) N-terminal truncated-Piwi remains bound to wildtype or nuclease-deficient Aub-CD in *Drosophila* ovaries. Immunoprecipitations were performed from ovaries of transgenic flies that express FLAG- and GFP-tagged proteins in germ cells under control of the MaG4 driver. FLAG-tagged proteins were immunoprecipitated in the presence (+) or absence (-) of RNase A followed by Western blotting to detect GFP- and FLAG-tagged proteins. The lower panel shows Input (Inp) and the unbound fractions (Unb) collected after immunoprecipitation.





***Figure 6. A model for assembly of Aub pre-cleavage complex that are recruit Piwi.***

This model is an elaboration of the model presented in Chapter 3. First, Aub loaded with a piRNA guide freely scans transcripts exiting the nucleus for complementarity. Upon recognition of a complementary target, Aub recruits components Qin and the Krimp-Ago3 complex. Prior to cleavage, Aub recruits factors SpnE and Mael, in addition to recruiting unloaded Piwi. The assembly of this complex is not yet understood, although SpnE and Mael might chaperone Piwi into the nucleus upon loading with piRNA.



***Supplemental Figure 1. Mean Pearson's correlation coefficient of reciprocal pairs of GFP-tagged or mK2-tagged Aub-WT, Aub-CD and Ago3-WT proteins co-expressed in nurse cells.***

Pearson's correlation coefficient quantifies the degree of co-localization of fluorophore-tagged protein pairs. Identical proteins tagged with either mK2 or GFP co-localize best. On average, Aub-CD and Ago3-WT localization correlates strongest of non-identical pairs, while Aub-WT and Ago3-WT correlate least well. Aub-WT and Aub-CD co-localization is on average correlated stronger than Aub-WT and Ago3-WT, but less well than Aub-CD and Ago3-WT. The mean Pearson's coefficient was obtained from eight independently measured nuclei expression, one mK2-tagged protein, and one GFP-tagged protein expressed using the same driver.

